



Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

Taking one for the team. Team reasoning and the possibility of collective alief.

Christensen, Bo Allesøe

Published in:

Book of abstracts. The Fifth Conference of the European Network on Social Ontology

Creative Commons License
Unspecified

Publication date:
2017

[Link to publication from Aalborg University](#)

Citation for published version (APA):

Christensen, B. A. (2017). Taking one for the team. Team reasoning and the possibility of collective alief. In *Book of abstracts. The Fifth Conference of the European Network on Social Ontology* (pp. 44-46)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.



LUND
UNIVERSITY



ENSO V

Book of Abstracts

The Fifth Conference of the European Network on Social Ontology

LUND | AUGUST 30TH – SEPTEMBER 1ST, 2017



Book of Abstracts

THE FIFTH CONFERENCE OF THE EUROPEAN NETWORK ON SOCIAL ONTOLOGY
LUND, AUGUST 30TH – SEPTEMBER 1ST, 2017



LUND
UNIVERSITY



ENSO V

About the conference

The European Network on Social Ontology was founded in 2007 with the goal to provide a platform for discussion and exchange of ideas, future research activities at European level and the organisation of a biennial international conference focusing on social ontology. Attendance to the conference allows participants to join the International Social Ontology Society (ISOS). The conference will be organised by the Metaphysics & Collectivity Research Group in cooperation with the Department of Philosophy at Lund University and The International Social Ontology Society.

ORGANISING COMMITTEE

Björn Petersson (Chair), Lund University; Gunnar Björnsson, Umeå University; Olle Blomberg, University of Copenhagen; Johan Brännmark, Malmö University; Åsa Burman, Stockholm University; Mattias Gunnemyr, Lund University; Tobias Hansson Wahlberg, Lund University; Megan Lambert, Lund University; Anna-Sofia Maurin, University of Gothenburg; Tomas Persson, Lund University; Jeroen Smid, Lund University; András Szígeti, Linköping University.

FUNDING

The organisers are grateful for generous financial support from the Swedish Research Council, Erik and Gurli Hultengren Fund for Philosophy at Lund University, Stiftelsen Roy och Maj Franzéns fond, and from the Department of Philosophy at Lund University.

THEMES

Possible themes include but are not limited to:

- I. The metaphysics of social reality
- II. Team reasoning, coordination problems, preference and value aggregation
- III. The nature of collective intentionality and agency

- IV. Moral and legal aspects of collective agency
- V. Social conventions, norms, and power
- VI. Non-human sociality

The Metaphysics of Social Reality

Social reality consists, ostensibly, of a diverse range of entities and phenomena, such as social events (e.g. elections), social properties (e.g. being a professor), social facts (e.g. Sweden's being a member of the EU) and social objects (e.g. corporations). Can we account for their reality and persistence in terms of entities and phenomena from other ontological categories or levels of reality, or are they best conceptualized as *sui generis* entities? What kinds of concepts – e.g. grounding, constitution, emergence, supervenience, reduction, etc. – are likely to be fruitful analytical tools in the endeavor of uncovering the metaphysics of social reality?

Team reasoning, coordination problems, preference and value aggregation

Human actions are, more often than not, interlinked. People frequently coordinate their actions, but just as often people aim for stronger forms of cooperation seeking to act together in order to accomplish tasks they could not solve in isolation. Many forms of such coordinated and cooperative activities involve the pooling of ideas, reasons, preferences and values among participants. However, these deliberative and epistemic aspects of joint and collective action raise well-known theoretical issues and practical difficulties. Among others, how successful are groups in aggregating their values and preferences? What are the forms and mechanisms of team reasoning? What is the relationship between collective intentionality and group rationality? We invite submission of abstracts relevant to these questions broadly conceived.

The nature of collective intentionality and agency

Individuals may arguably jointly attend to an object, share emotions, intend to act together, share beliefs about the world, and collectively endorse common values. But in which sense can attitudes be shared or collective? Can such attitudes be ascribed to groups as such or are there other ways of distinguishing collective from individual intentionality? Which conceptual resources are required for collective acting and intending? How is it related to self-awareness, and to social cognition? The nature of collective intentionality and agency is essential to our understanding of the capacities and limitations of groups and institutions, and for questions about their moral and legal status.

Moral and legal aspects of collective agency

Moral claims often concern groups of agents or organisations such as states or corporations whose activities centrally involve such groups. These “collective” entities are said to have obligations and are held responsible for actions and outcomes. But can they really be moral agents? If not, can they really have obligations or bear responsibilities? If they can, what is the relation between their obligations and responsibility and that of members of the groups? Similar questions arise in the legal realm, where aspects of complicity in collective wrongdoing and duties to redress collectively brought-about harms have received particular attention.

Social conventions, norms, and power

Conventions and conventional understandings play a normative role in society, regulating interactions and helping people to coordinate their behavior. But by doing so they also inevitably privilege, whether explicitly or effectively, certain social positions above others, e.g., through constructions of normality. These issues are often addressed in gender studies, critical race theory, disability studies, and other approaches, as well in recent philosophical discussions about epistemic injustice, but can also be addressed through an understanding of the ontology of conventions, norms, and power.

Non-human sociality

Collective entities and phenomena in our societies exist as solutions to human needs. But both human needs, and their solutions, are dependent on biologically evolved substrates. Sociality as such is an integrated part of most life forms. Collectivity in a comparative perspective places human sociality in an evolutionary context where similarities and differences between species guide theories about function and deep history of human collectivity, and what mechanisms are needed, or involved, for specific types of collectivity. Such comparisons need not be limited to animals, but also other forms of nonhumans, such as robots with artificial intelligence, can contribute to modelling the causes of human collectivity.

Contents

About the conference	3
Organising committee	3
Funding	3
Themes	3
Contents	7
Book Symposia	13
Francesco Guala, <i>Understanding Institutions: The Science and Philosophy of Living Together</i> , Princeton U P 2016	14
Participants:	
Rico Hauswald	
Elisabeth Pacherie	
Wlodek Rabinowicz	
Francesco Guala	
Institution Types and Institution Tokens: An Unproblematic Distinction?	
- Commentary on Francesco Guala's <i>Understanding Institutions</i>	15
Rico Hauswald	
Kirk Ludwig, <i>From Individual to Plural Agency: Collective Action 1</i> , Oxford U P 2016	16
Participants:	
Sara Rachel Chant	
Michael Schmitz	
Raimo Tuomela	
Kirk Ludwig	
From group minds to group mindedness	
- Commentary on Kirk Ludwig's <i>From Individual to Plural Agency</i>	17
Michael Schmitz	

Comments on Kirk Ludwig's book <i>From Individual to Plural Agency</i> , 2016, OUP <i>Raimo Tuomela</i>	18
Keynotes	20
<i>Stephen Butterfill</i>	
<i>Sara Rachel Chant</i>	
<i>Christian List</i>	
<i>Karen McComb</i>	
<i>Julie Zahle</i>	
From Minimal to Complex Collective Actions <i>Sara Rachel Chant</i>	21
Levels: Descriptive, Explanatory, and Ontological <i>Christian List</i>	22
Sociality, communication and cognition in non-human animals <i>Karen McComb</i>	23
Explanatory and Ontological Levels in the Methodological Individualism-Holism Debate <i>Julie Zahle</i>	24
Presentations	25
Social Cognition, Cooperation, and Hypothetical Contracts <i>Matteo Bianchin</i>	26
Moral demands on groups <i>Gunnar Björnsson</i>	29
More (unsound) arguments for the common knowledge condition <i>Olle Blomberg</i>	31
Minimal approaches to joint action <i>Ingar Brinck</i>	33
Money as an Interactive and Subjective Social Kind <i>Eyja Brynjarsdóttir</i>	36
The status function account of human rights – Too institutional, or not institutional enough? <i>Åsa Burman</i>	38
Normative Representations and Emotions <i>Antonella Carassa, Marco Colombetti</i>	40
Ghost Stories. Toward an Hegelian Approach to Social Ontology <i>Alessandro De Cesaris</i>	42
Taking one for the team. Team-reasoning and the possibility of collective alief <i>Bo Allesøe Christensen</i>	44
When the (Democratic) State Acts, does its Citizenry? <i>Stephanie Collins and Holly Lawford-Smith</i>	46
Collective reasoning and collective responsibility gaps <i>Hein Duijf</i>	48

Causal Responsibility for Emissions in Global Supply Chains <i>Anton Eriksson</i>	51
Social Cognition, Empathy, and Agent-Specificities in Cooperation <i>Anika Fiebich</i>	52
Theories of Group Agency: Preliminary Experimental Data <i>Brian Flanagan</i>	54
Organizations: an Ontological Approach <i>Francesco Franda</i>	55
Towards a Social Ontology of Work <i>Rachel Fraser</i>	56
A New Method for Social Preference Aggregation <i>Kalle Grill</i>	58
Explaining Moral Responsibility for Structurally Produced Harms <i>Mattias Gunnemyr</i>	60
Collective Moral Agency and Identity Over Time <i>Niels de Haan</i>	63
Causal Powers and Social Ontology <i>Tobias Hansson Wahlberg</i>	65
What's the Difference Between Money and Gender? Social Construction, Critique, and Change <i>Frank Hindriks</i>	66
How Do Non-Joint Commitments Come Into Being?: An Attempt at Cultural Naturalism <i>Ingvar Johansson</i>	67
The Explanatory Power of the Axiological Objects of Collective Intention <i>Taina Kalliokoski</i>	68
Declarations and institutional reality <i>Arto Laitinen</i>	70
Social Construction and Universality <i>Esa Diaz Leon</i>	71
How to Think About Constructions: Towards an Account of the Acquisition of Social Kind Concepts <i>Robin Guid Loehr</i>	73
Heuristics, bounded rationality, and joint action <i>Judith Martens</i>	76
Collective memory: Metaphor or reality? <i>Kourken Michaelian</i>	78
A theory of collective self-awareness <i>Anna Moltchanova</i>	81
Collective Intentionality as the common ground of normative experiences <i>Gloria Mähringer</i>	83
Institutions as Shared Plans: Practical Reasoning in I and We Modes <i>Pekka Mäkelä, Raul Hakli and S.M. Amadae</i>	85

Objects, People, and Powers: The Deontic Theory of Money <i>Benjamin Neeser</i>	86
Nostra Res Agitur? Programmatic Contexts and The Disentangled We <i>Henning Nörenberg</i>	88
Social reality without levels <i>Gianluca Pozzoni</i>	90
Macro phenomena without macro manifestations – why the multiple realization argument is incompatible with social science <i>Gustav Ramström</i>	92
Collective Agency and Individual Criminal Responsibility <i>Eirik Julius Risberg</i>	94
The Practical Significance of Group Reasons for Individual Agents <i>Abe Roth</i>	96
Discrimination and Collaboration in Social Networks <i>Hannah Rubin and Cailin O'Connor</i>	98
Money and Deontic Power <i>Joshua Rust</i>	100
From Coordination to Cooperation. Switching to Team Reasoning <i>Jules Salomone</i>	102
Non-observational knowledge and Joint action <i>Glenda Lucila Satne</i>	104
Collective Responsibilities of Random Collections: Plural Self-Awareness Among Strangers <i>Hans Bernhard Schmid</i>	106
Co-subjective consciousness creates collectives <i>Michael Schmitz</i>	108
The non-epistemic origins of research strongholds <i>Mike D Schneider</i>	110
Joint and mutual obligations <i>David Schweikard</i>	112
Grounded Composition? <i>Jeroen Smid</i>	114
Artificial agents in the realm of social cognition <i>Anna Strasser</i>	115
Nested Groups: Membership and Parthood <i>David Strohmaier</i>	117
Collective Hatred and Hatred towards Collectives <i>Thomas Szanto</i>	119
Exploitation and Joint Action <i>András Szígeti and Erik Malmqvist</i>	120
Is Art an Institution? Are Artworks Social Objects? Does the Philosophy of Art Fit into Social Ontology? <i>Enrico Terrone</i>	122

Two Ways of Feeling Together <i>Gerhard Thonhauser</i>	124
Collectivity in Blockchain based Environments <i>Sabine Thuermel</i>	126
Collective testimony and second personal commitment <i>Leo Townsend</i>	127
On the moral normativity of socially constructed norms <i>Laura Valentini</i>	129
The Plurality in a Plural Subject: The Phenomena of Dissenting and Dissidence as Communal Phenomena <i>Xiaoxi Wu</i>	131
What would you like me to drink? The Epistemic Foundations of Salience-based Coordination <i>Vojtěch Zachník</i>	134
Belief-Revision, Epistemic Contribution, and Polynormativity <i>Marko-Luka Zubčić</i>	136

Maps	137
------	-----

Program	141
Day 1 (Wednesday Aug 30 th)	142
Day 2 (Thursday Aug 31 st)	143
Day 3 (Friday Sep 1 st)	145

Book Symposia

Francesco Guala, *Understanding Institutions: The Science and Philosophy of Living Together*, Princeton U P 2016

PARTICIPANTS:

RICO HAUSWALD

Institution Types and Institution Tokens: An Unproblematic Distinction?

Technische Universität Dresden

ELISABETH PACHERIE

Solution thinking and team reasoning: how different are they?

Institut Jean Nicod, CNRS

WLODEK RABINOWICZ

Lund University

FRANCESCO GUALA

Reply

University of Milano

Institution Types and Institution Tokens: An Unproblematic Distinction?

- Commentary on Francesco Guala's *Understanding Institutions*

RICO HAUSWALD

Technische Universität Dresden

ABSTRACT

The distinction between institution types and institution tokens plays an important role in Francesco Guala's philosophy of institutions. In this commentary, I argue that this distinction faces a number of difficulties that are not sufficiently addressed in *Understanding Institutions*. In particular, I critically discuss Guala's comparison between the taxonomy of organisms and the taxonomy of institutions, consider the semantics of institution terms on different levels in this taxonomy, and argue for an alternative solution to the problem of how to reconcile reformism and realism about institutions like marriage.

Kirk Ludwig, *From Individual to Plural Agency: Collective Action 1*, Oxford U P 2016

PARTICIPANTS:

SARA RACHEL CHANT

Ludwig's Universalism About Collective Action

Tulane University

MICHAEL SCHMITZ

From group minds to group mindedness

University of Vienna

RAIMO TUOMELA

*Comments on Kirk Ludwig's book *From Individual to Plural Agency*, 2016, OUP*

University of Helsinki and University of Munich

KIRK LUDWIG

Reply

University of Indiana

From group minds to group mindedness

- Commentary on Kirk Ludwig's *From Individual to Plural Agency*

MICHAEL SCHMITZ
University of Vienna

ABSTRACT

The philosophy of collectivity is still often driven by fear of group minds. Defenders of content accounts like Michael Bratman attempt to analyze away all occurrences of “we” as a subject of intentions, proposing that at least small-scale collectivity can be entirely explained in terms of attitudes of the form “I intend that we J”. Proponents of mode accounts like John Searle reject this and embrace group mindedness as irreducible, but are also dismissive of collective subjects. In his recent book, Kirk Ludwig makes valid criticisms of the mode account and of Margaret Gilbert’s version of a plural subject account and develops a very rich and improved version of the content account. However, I will argue that this analysis is still inadequate and subject to counterexamples, and that it throws out the baby of the group mindedness essential to understanding joint intention with the bathwater of group minds. I conclude by sketching an account of groups as subjects of intentions that responds to Ludwig’s criticisms of Searle and Gilbert. It explains how group mindedness ties the group members together so that they can jointly be the subjects of intentions and other attitudes. A group has minds, not a mind. As a subject of intentions it is therefore essentially a plural entity, while both the notion of a group mind and the content account try to reduce it to something singular. Joint commitment is one form of group mindedness that ties subjects together into collectives, but there are also others such as e.g. joint attention, joint skills, or joint habits.

Comments on Kirk Ludwig's book *From Individual to Plural Agency*, 2016, OUP

RAIMO TUOMELA

University of Helsinki and University of Munich

ABSTRACT

The main ideas of the book

It is rich in content and extensive in scope. There are useful summaries in both parts of the book. On the whole, the book is an exceptionally well-argued, high-level contribution to analytic single-agent and plural-agent action theory.

Social groups are viewed as plural agents in the book and they thus ontologically consist of a collection of single agents who generally are capable of shared we-intending that they act together in specific ways.

Ludwig's book develops an analysis of individual action, agency, and intention, with special focus on the logical form of singular action sentences. The logical form is characterized basically in terms of refined and extended Davidsonian event-based analysis.

In Ludwig's account the action verb is treated as introducing two quantifiers over events. To say that an agent melted the chocolate is to say roughly that he was the immediate agent of some event that brought it about that the chocolate melted (that an event of melting of the chocolate occurred).

The second part (of this book) develops an account of collective action and intention. It focuses on an account of the logical form of plural action sentences ("We built a bridge") and on sentences attributing "we-intentions" to action (e.g. "We intended to build the bridge") as well as on the content of individuals' intentions (viz. "we-intentions") when they are engaged in joint intentional action.

Plural action sentences are typically ambiguous between a *distributive* and a *collective* reading, where the ambiguity is due to quantifier scope ambiguity but not lexical ambiguity concerning intending. Singing the national anthem could involve that we each did so (distributive reading), or that we did so together (collective reading). On the distributive reading, for each of us there is an event of which he is the sole agent and it is the singing of the national anthem. The collective reading is obtained by reversing the order of the quantifiers. On it, there is an event, viz. a singing of the national anthem, of which each of us is an agent. Here group agency is taken by Ludwig to involve multiple agents of a single event, not an event of which there is a single *group* agent. In both readings the *we* is present. Only the group members are agents here.

For Ludwig collective intentional action is the most fundamental form of social reality. E.g. social institutions, practices, and interaction depend on it. Our foundational understanding of social reality is said to depend on understanding the nature of collective intentional behavior and how it differs from individual intentional behavior. An account of collective intentional behavior should be built on an understanding of individual intentional behavior, and the account should be built on a proper understanding of the logical form of the relevant descriptive

sentences. This “rests on the assumption that the logical and conceptual resources of our discourse about the social are fitted to the phenomena we use it to talk about”. Laying bare the logical form of our discourse tells us the fundamental ontology of the domain we are interested in. This view is somewhat problematic as it seems to depend on our *a priori* common sense ideas rather than on empirical research.

As to we-intentions, consider again the case of singing the national anthem intentionally. On the distributive reading, the members of “we” sing the anthem separately or severally while in the collective case they sing the anthem jointly so that each of them nevertheless is in the agency relation to the end result of the anthem having been sung. I agree with Ludwig’s account about this in principle but argue that the plural subject group can be the agent of anthem singing in an approximate functional sense of acting (see Tuomela, 2013, ch. 2).

For Ludwig the plural subject group is not an agent, only its members are agents when acting intentionally on their relevant we-intentions. A we-intention is an intention of an individual that is directed toward an action of his while at the same time purporting to contribute to a collective action in which that individual has a part, which itself results from everyone having the same collective action plan associated with his intention when he acts as part of the group.

A central achievement of Ludwig’s account is that it gives detailed and innovative technical representations of many central notions and phenomena that action theory is about, although this takes place in individualistic terms and without the assumption of a *sui generis* “we-mode” or the like notions. The book purports to show that collective action can be appropriately analyzed in terms of individualistic notions.

We-intention is a central element in Ludwig’s account (see chapter 13) and also in a somewhat different sense in my own account that in contrast to his content-based account uses “thick”, *we-mode* we-intentions (e.g. in my 2013 book). It does not agree with Ludwig’s thesis that we-intentions should be analyzed solely on the basis of concepts which are already in play in understanding individual intentional action. The we-mode is with us in many cases of collective thinking and acting – people are typically cooperative and also disposed to act together in the we-mode. (Some experimental work by e.g. Colman supports this claim.)

I have argued that we-mode we-intentions are in many cases (e.g. Paretian collective action dilemmas) needed for social theorizing and for the best description and explanation of the social world. My account centrally involves the “adverbial” view given on pp. 36-38 of Tuomela (2013) concerned with somebody intending or believing (etc.) a content or acting on it. These activities can be characterized partly in terms of the phenomenal features of the intender’s awareness of the group as subject (and his experiencing others as co-subjects) when intending (etc.) and also in terms of what special overt “groupish” features the satisfaction of the intention involves. In the case of the I-mode we are dealing with individualistic activities of the kind in which a single agent intends as a single person in contrast to intending in the we-mode as a group member.

Keynotes

STEPHEN BUTTERFILL

Cooperation and Motor Representation

University of Warwick

SARA RACHEL CHANT

From Minimal to Complex Collective Actions

Tulane University

CHRISTIAN LIST

Levels: Descriptive, Explanatory, and Ontological

The London School of Economics and Political Science

KAREN MCCOMB

Sociality, communication and cognition in non-human animals

University of Sussex

JULIE ZAHLE

Explanatory and Ontological Levels in the Individualism-Holism Debate

University of Bergen

From Minimal to Complex Collective Actions

SARA RACHEL CHANT

ABSTRACT

Collective action theory models both the salient questions and answers for collectives after individual action theory. Since the first problem for individual action theory concerns the conditions for something's being an individual action, the first problem for collective action theory concerns the conditions for something's being a collective action. Likewise, since standard answers to the first problem in individual action concern the mental states of an agent (e.g. their beliefs, desires, plans and intentions), standard answers to the first problem in collective action theory concern the mental states of the group. Because this approach is straightforwardly applied from individuals, and individuals are saliently different from groups, providing a unified analysis across the wide range of actions taken by groups proves far more challenging than its individual counterpart. As a result, most theories of collective action focus on 'small-scale' actions to discuss the minimal conditions for something's being a collective action, with the hope that they can extend that minimalist account to cover the very wide array of complex actions of groups. In this talk, I will suggest that the standard minimalist proposals are not minimal enough, and that that mistake arises from the general approach taken in collective action theory. Once we recognize that mistake and consider truly minimal collective actions, we may begin to see the way toward a unified theory from minimal to complex collective actions.

Levels: Descriptive, Explanatory, and Ontological

CHRISTIAN LIST

The London School of Economics and Political Science

ABSTRACT

Scientists and philosophers frequently speak about levels of description, levels of explanation, and ontological levels. The aim of this talk is to present a unified framework for modelling levels and to illustrate its relevance to social ontology. I give a general definition of a *system of levels* and show that it can accommodate descriptive, explanatory, and ontological notions of levels. I further apply the framework to some salient philosophical questions: (1) Is there a linear hierarchy of levels, with a fundamental level at the bottom? (2) Are there emergent properties, such as emergent indeterminism and randomness in social systems? (3) Are high-level descriptions (such as “macro” descriptions) reducible to lower-level ones (such as “micro” descriptions)? Although I use the terminology of “levels”, the proposed framework can also represent “scales”, “domains”, or “subject matters”, where these are not linearly but only partially ordered by relations of supervenience or inclusion.

Sociality, communication and cognition in non-human animals

KAREN MCCOMB

University of Sussex

ABSTRACT

Communication through vocal, visual and olfactory signals is a fundamental mechanism at the basis of animal and human social systems. In my talk I will explain how such signals, when presented to animals in naturalistic experiments (in isolated modalities or in combination), can be used to investigate the cognitive abilities that underlie complex social systems and reveal how animals interpret their social and emotional worlds. I will focus on four areas that are highly relevant to exploring origins in animals of advanced human traits:

- Numerical assessment
- Social knowledge
- Cross-modal recognition and representational knowledge
- Emotional communication

In considering these areas I will draw from my research on lions, elephants and horses and illustrate how it is possible to delve deeply into the mental lives of social animals using naturalistic paradigms. Such investigations have revealed a range of advanced cognitive abilities and high levels of social and emotional understanding in these non-verbal species that had hitherto been considered uniquely human.

Explanatory and Ontological Levels in the Methodological Individualism-Holism Debate

JULIE ZAHLE

University of Bergen

ABSTRACT

The methodological individualism-holism debate is commonly formulated in level terms. Accordingly, methodological individualists are said to defend that view that social scientists should only offer individual-level explanations, whereas methodological holists are taken to maintain that social scientists should sometimes advance individual-level explanations; other times social-level explanations.

In this paper, I argue that the level-framework has only limited applicability to discussions of methodological individualism and holism: relative to supervenience, emergence, and other accounts of the social world as leveled, widely endorsed forms of methodological individualism simply resist formulation in individual-level terms. This point has largely gone unnoticed yet as I also show, it has a number of important implications for the ongoing methodological individualism-holism debate.

Presentations

Social Cognition, Cooperation, and Hypothetical Contracts

MATTEO BIANCHIN

Università di Milano-Bicocca

ABSTRACT

Social contract theory – in the hypothetical contract version – traditionally faces three objections. That hypothetical contracts are not binding (Dworkin 1973); that they model actors according to an unrealistically abstract conception of self and agency (Sandel 1983); that the theory is either circular or redundant (Nagel 1973; Pettit 1993). I argue that current work in social cognition provides the tools to cope with the second and the third worry, and to deal with the first once it is framed as a stability problem. I draw on Tomasello's work on the psychological infrastructure of cooperation and on Goldman's simulation theory to figure out the mechanisms at work, and argue that they provide agents with a framework for handling hypothetical contracts, grasping their normative content, supporting the motivation to comply. The background idea is that social contract theory can be given a plausible mechanistic reading. I frame the issue in the context of Rawls' understanding of the original position as a "device of representation" designed to convert a question of justification into a deliberative problem. The original position models the conditions under which agents that conceive themselves as free and equals rational beings are supposed to reach an agreement and thus conveys an impartial and fair point of view that constrains what can be put forward as a good reason in deliberating the principles of justice. beings (Rawls 1971, 1985; Freeman 2007; Barry 1995). Critics objected that this misconstrues how actual agents understand themselves and circularly introduces in the initial situation the normative contents that are supposed to be the output of the contract. Current research on social cognition, however, arguably supplies the tools to work out a psychologically feasible reading. Drawing on Tomasello's work on the cognitive infrastructure of cooperation I take that (a) cooperative activities require understanding the equivalence between self and other and therefore a capacity for social cognition that supports perspective taking and role reversal; and (b) the pro-social motives stemming from the early inclinations to help and share develop through social interaction into a disposition to reciprocity (Tomasello 2009, 2014). I further maintain that perspective taking and role reversal are supported by simulative mindreading and rest on a single mechanism of imaginative self-projection that works both in intrapersonal action planning and in interpersonal cooperation (Goldman 2006, 2013; Bruckner, Carroll, 2007).

Based on this, I argue that, although the limited mindreading capacities and pro-social motives involved in early joint actions do not support the Kantian interpretation of the original position suggested by Rawls, a suitably abstract understanding of agency can be taken to develop out of them as a representational theory of mind is acquired in connection with the syntax of sentential complements and social cognition accordingly rearranges in line with a propositional attitude folk psychology (Rakoczy 2015; de Villier 2007). In this context agents generalize the conditions under which joint actions are performed and learn to cooperate in

anonymous settings (Tomasello, Rakoczy 2003). We should expect that the self-other equivalence that goes along with early joint actions also generalizes to the effect that individuals come to conceive themselves mutually as rational agents endowed with a capacity for self-monitoring and self-governance.

I contend that the fair and impartial point of view captured by the original position becomes accessible as agents come to see themselves as free and equal in this minimal sense and consequently acquire a capacity to reason and act according to the normative expectations of a "generalized other". A related shift can be expected to occur in their understanding of justice, to the effect that the early aversion to unequal distributions (Ulber, Hamann, Tomasello 2017) develops into a full-blown conception of justice as fairness.

Under this reading social contract theory does not look circular, as it rests on the working of socio-cognitive mechanisms that are not morally loaded. In particular, no moral assumption is build in the understanding of agency and in the self-conception of agents that underlay the original position, in contrast with the idea of moral persons allegedly found in the political culture of democratic societies (Rawls 1993). Deliberating in the original position just results from recruiting the self-other equivalence that goes along with early joint actions under the abstract conception of agency individuals acquire in the course of developing their socio-cognitive skills.

Finally, I conjecture that the early disposition to reciprocate may be recruited to yield a general system of mutual normative expectations as social cognition develops along the line sketched above. This may account for the motivation to agree on fair terms of cooperation and to comply with them once they are in place, which allows a fair system of cooperation to be stable or self-supporting in Rawls' sense.

REFERENCES

- Bruckner, R., Carroll, D. (2007). Self-Projection and the Brain. *Trends in Cognitive Sciences* (11) 2: 49-57.
- De Villiers (2007). The Interface of Language and Theory of Mind. *Lingua* (117) 1, 1858-1878.
- Dworkin, R. (1973). The Original Position. *The University of Chicago Law Review* (40) 3: 500-533.
- Enoch, D. (2015). Against Public Reason. In D. Sobel, P. Vallentyncae, S. Wall (eds.), *Oxford Studies in Political Philosophy*. Oxford OUP.
- Fenici, M. (2017), What is the Role of Experience in Children's Success in the False Belief Test: Maturation, Facilitation, Attunement or Induction?. *Mind & Language* 32(3): 308-37.
- Freeman, S. (2007). *Justice and the Social Contract*. Oxford: OUP.
- Gaus, G., Thrasher, J. (2015), Rational Choice and the Original Position:. In T. Hinton (ed.). *The Original Position*. Cambridge. Cambridge University Press.
- Goldman, A. (2006). *Simulating Minds*. Oxford: OUP.
- (2013). *Joint Ventures*. Oxford: OUP.
- Nagel, Th. (1973). Rawls on Justice. *The Philosophical Review*, (82) 2: 220-34.
- Pettit, Ph. (1993). *The Common Mind*. Oxford: OUP.
- Rakoczy, H. (2017). In Defense of a Developmental Dogma: Children Acquire Propositional

- Attitude Folk Psychology around Age 4. *Synthese* 194(3): 689-707.
- Rawls, J. (1971). *A Theory of Justice*. Cambridge (MA): Harvard University Press.
- (1993). *Political Liberalism*. New York: Columbia University Press.
- Sandel, M. (1983). *Liberalism and the Limits of Justice*. Cambridge: CUP.
- Tomasello, M. (2009). *Why We Cooperate*. Cambridge (MA): MIT Press
- (2014). *A Natural History of Human Thinking*. Cambridge (MA): Harvard University Press.
- (2016), *A Natural History of Human Morality*. Cambridge (MA): Harvard University Press.
- Tomasello, M., Rackoczy H. (2003). What Makes Human Cognition Unique? *Mind & Language* (18) 2: 121-147.
- Ulber, J, Hamann, K., Tomasello, M. (2017). Young children, but not chimpanzees, are averse to disadvantageous and advantageous inequities. *Journal of Experimental Child Psychology* (155): 48–66.

Moral demands on groups

GUNNAR BJÖRNSSON

Stockholm University

ABSTRACT

Moral demands, it seems, can only fittingly be directed at entities capable of recognizing and being guided by such demands, including by holding themselves responsible for failures to comply. Such capacities are hard-won, and moral demands do not typically target animals, young children, or adults with certain limited cognitive or volitional capacities. They might perhaps be satisfied by some entities whose thinking and actions are constituted by the thinking and actions of groups of individuals (Björnsson and Hess 2017; cf French 1984; Hess 2014; List and Pettit 2011; Rovane 1998; 2014a; b). But most groups of individuals seem to fall far short with respect to internal organization to be target of any moral demands going beyond those directed at the individuals (cf. Collins 2013; Lawford-Smith 2015).

In spite of this, I have argued elsewhere, individuals in unstructured groups can *share* obligations, and be *jointly* blameworthy for failures to live up to these (Björnsson 2011; 2014; Forthcoming; see also, among many others, Kutz 2000; May 1992; Schwenkenbecher 2014; Wringer 2010; 2016). The key to this argument is that although responsibility and obligations generally depend on what can be fundamentally demanded of individual moral agents, they do not correspond directly to such demands. Fundamental moral demands are demands to care about morally important matters, where to care about something in the relevant sense is to be disposed to notice what promotes or threatens the object of caring and to be motivated to promote or protect it accordingly. For blameworthiness and one closely related kind of all-out moral obligation or requirement, I have then defended the following claims:

BLAMEWORTHINESS: X is morally to blame for Y if and only if Y is morally bad and explained in a normal way by X's substandard caring (Björnsson 2011; 2017a; b; Björnsson and Persson 2012).

REQUIREMENTS AS ENSURED BY CARING (REC): X is morally required to ϕ if and only if X's not ϕ ing would be morally bad and X's ϕ ing would be ensured, in a normal way, by X's caring satisfying the standards (Björnsson 2014; Björnsson and Brülde 2017).

Though demands on caring only apply to moral agents, something morally bad might be the normal upshot of the substandard caring of a number of individuals, as when a lack of appropriate concern for the environment among a great number of individuals leads or would lead to the destruction of ecosystems. In such a case, the individuals might together be blameworthy for the destruction, even though no one individual could have made a difference. Similarly, it is often the case that good outcomes would be ensured in a normal way if a number of individuals cared to the extent that can be reasonably demanded of them. In such cases, they might share an obligation to bring about this outcome, even if no one individual could make a

difference. On the proposed view, shared blameworthiness and obligations cannot be reduced to individual counterparts, as there are cases where no individual caring explains or ensures a certain outcome, but where the states of caring of all members of a group, taken together, do. In this sense, shared blameworthiness and obligations are holistic.

In this talk, I argue that shared blameworthiness and obligations are subject to another level of holism, with respect to the moral demands on individual caring on which they rest. The reasonableness of such demands, I argue, depend largely on how hard it is for an agent to achieve and uphold various levels of caring or, differently put, on what levels of caring he could achieve or uphold given that various amounts of cognitive and volitional resources was put into this. Importantly, degrees of resources required depends on the sort of feedback and support the agent can receive. Such feedback can be intrapersonal, and can be broadly understood as the operation of the individual's conscience. Because of this, capacities for caring depend on the capacity for conscience. But feedback can also be interpersonal, as when others react to the agent's failures or successes in relation to ideals of caring, or to the failures and successes of others, helping the agent identify the right levels of caring and be motivated to achieve these.

Notably, the availability of the latter kind of support will depend on whether others care about the objects in question and are concerned with getting levels of caring right. If they do, then the individual might have a much easier time achieving and upholding better levels of caring. This, in turn, means that when we ask what how hard it would be for a group of individuals to achieve and uphold a certain level of caring, the answer to that question might be different than if we ask how hard it would be for an average individual to achieve that level. In asking the latter question, we need to take into account the actual levels of caring in her social circumstances. In asking the former question, we are in effect assuming that all are putting efforts into their caring. Given that the group is a social connected, that means that their efforts could be mutually supporting, making caring at that level considerably easier. As a consequence, what levels of caring can be reasonably demanded of a group of individual agents is not a function what can be demanded of each individually.

A central upshot of the argument is that under certain circumstances, a group of people can share blame for something (destroying an ecosystem, upholding structural injustice) even though all members individually cares as can be reasonably demanded of them given the actual levels of caring of the others. Similarly, the group can have an obligation to do something (leaving the ecosystem intact, eradicate injustice) even though the group would not do this even if each individual cared as can be reasonably demanded of him given the others' actual level of caring.

More (unsound) arguments for the common knowledge condition

OLLE BLOMBERG

Lund University

ABSTRACT

Most reductionist accounts of joint intentional action, or “shared intention”, include a condition that the intentions and other attitudes of the participants that make up their shared intention must be common knowledge between them. Some authors have argued against the common knowledge requirement (see e.g. Kutz 2000; Blomberg 2016; Ludwig 2016). In this talk, I consider but ultimately reject two potential new arguments for keeping this requirement. Neither of the arguments has been articulated in print, but several philosophers have tentatively suggested the first in conversation, and the second at least seems to be implicit in (Rovane 1998, ch. 4). There are many ways in which one might understand common knowledge, mutual knowledge, or metaphorically speaking, “openness”. One can understand it in an undemanding way simply as a “negatively characterised notion [of] mutual absence of doubt [or false belief]” (Davies 1987), or as a more demanding interpersonal state of affairs from which an infinite hierarchy of states of awareness could be derived (see e.g. Lewis 1969).

The first new argument is relevant to the former undemanding understanding of common knowledge—or rather, of openness. According to this argument, cases in which one or more participants have a false higher-order belief regarding a participant’s intention (for one such case, see Blomberg 2015) will involve some kind of deviant causation. So while openness might not be required for shared intention per se, it is required for shared intentions to non-deviantly lead to joint intentional action. In response, I show that cases involving such false higher-order beliefs need not involve any antecedential (basic), consequential (secondary) or tertiary deviance or waywardness with respect to the individual participatory intentions that are building blocks of the shared intention, nor with respect to the shared intention as such (see Mele 1987 for the tripartite classification of types of waywardness). I also consider whether the false higher-order belief could undermine a joint action’s status of being jointly intentional due to evidential requirements on intentional action (see condition (iii) of Mele and Moser’s (1994) account). Here, I argue that the false higher-order beliefs would only undermine this status if they sufficiently significantly compromise the participants’ control over the action. And the false higher-order beliefs need not have this compromising effect.

The second argument has the following starting point: Joint intentional action and (most) singular intentional actions are species of one and the same phenomenon. They are species of complex intentional action, that is, larger actions composed of many smaller intentional actions (think of cooking dinner on your own and cooking dinner together with a friend). Now, Rovane argues that the psychosocial conditions of joint intentional action perfectly parallel those of what she calls “long-term activities”. The former requires that “the participants must have mutual knowledge of one another’s thoughts and actions”, while the latter requires that the agent can remember its past intending and anticipate the continuing of its intending as well as

that the agent can take it for granted that it has this memory and anticipation (and presumably, that it can take it for granted that it can take this for granted, and so on and so forth) (Rovane 1998, p. 145). This could be made into an argument for the common knowledge condition on shared intention: Joint intentional actions are composed of smaller intentional actions that are unified in the same way that complex singular intentional actions are unified across time within a single agent; since complex singular intentional action requires a kind of openness or common knowledge between time-slices of the agent, common knowledge should therefore similarly be required between participants in joint intentional action. While I endorse the starting point of this argument, I argue that the second premise is false. Complex singular intentional action does not require the openness or common knowledge between time-slices of an agent that would be necessary for the argument to deliver the sought conclusion.

REFERENCES

- Blomberg, O. (2016). Common Knowledge and Reductionism about Shared Agency. *Australasian Journal of Philosophy*, 94(2), 315–326.
- Davies, M. (1987). Relevance and mutual knowledge. *Behavioral and Brain Sciences*, 10, 716–717.
- Kutz, C. (2000). Acting Together. *Philosophy and Phenomenological Research*, 61(1), 1–31.
- Lewis, D. 1969. *Convention: A Philosophical Study*, Cambridge, MA: Harvard University Press.
- Ludwig, K. (2016), *From Individual to Plural Agency: Collective Action: Volume 1*, Oxford University Press.
- Mele, A. R. (1987). Intentional action and wayward causal chains: The problem of tertiary waywardness. *Philosophical Studies*, 51(1), 55–60.
- Mele, A. R., & Moser, P. K. (1994). Intentional action. *Nous*, 28(1), 39–68.
- Rovane, C. (1998), *The Bounds of Agency: An Essay in Revisionary Metaphysics*, Princeton University Press, Princeton, New Jersey.

Minimal approaches to joint action

INGAR BRINCK
Lund University

ABSTRACT

A joint action is such that could not have been brought about by a single agent, nor by a group of agents acting individually or in parallel. Jointness implies that an action is done together in the sense that each agent has the intention to perform the action in question, each acts so as to contribute to achieving the action, and each adjusts his or her individual contributions to those of the other agents to successfully perform the action with the others. An additional condition is that the agents share the goal of the action. This conception of joint action raises questions about what it means to share intentions and goals, what an agent has to do or know in order to adjust his or her own actions to the others' actions, etcetera. Rational accounts of joint action tend to assume that joint action requires higher-order representation, higher-order intention, propositional thought, and inferential reasoning. Such accounts have been criticized for being too strong, for presupposing capacities for logical reasoning, meta-representation, and mind-reading that are not in fact necessary for many forms of joint action, and for demanding forms of commitment that rarely are justifiable in everyday life. They also have been deemed implausible for large-scale joint action that involves a great number of participants such as performed by a corporation, state, or army. My talk will take its starting-point in the former kind of criticism.

Attempts to deal with the criticism by presenting dual or hybrid accounts tend to either suffer from making similar assumptions about the nature of joint action as the original accounts, focusing on the type of mental or inner states that supposedly are required for joint action and on how agents think about each other as a group or collective with regard to the goal of action; or they fail to meet the criticism in keeping to the rational analysis while adding mechanisms or routines to the account meant to (a) substitute for rational thought concerning less complex forms of joint action or forms that belong to the standard repertoire in everyday life, and sometimes (b) explain how children and animals can engage in actions that involve several agents and a common goal.

I will sketch an alternative response to the criticism that conceives of joint action as embodied, based in face-face interaction between emotionally engaged and attentive agents and relying on artefacts in the local environment to scaffold performance. The following points are all relevant for developing an alternative to the rational account and I will address a few of them in my talk:

- I. Most of the notions that turn up in rational accounts, e.g. common ground, mutual knowledge, commitment, can be cashed out in terms that do not require logical reasoning, recursion, or higher-order representations.
- II. Research in neuroscience on e.g. mirror neurons, social intention, coordination and entrainment suggest that human beings are biologically prepared for group or joint actions.
- III. Prima facie there is no reason to model joint on individual action, because dedicated

(evolved) mechanisms are likely to be distinct, and both individual and joint actions are distributed and rely on a great variety of different scaffolding.

- IV. Some problems that have been raised against non-rational accounts are relevant merely for motor-only theories and dual (rational-plus-motor) theories, which attempt to explain intentional actions by processes in the brain that are automatic, implicit and cognitively impenetrable. Plural or hybrid accounts invoking pre-reflective processes that are open to monitoring and control are not hit by this criticism.
- iv. Non-rational accounts raise questions about the normative relations between agents concerning e.g. trust, commitment, obligation, and responsibility, and about cognitive abilities such as making predictions about own and others' behaviour and outcomes of actions and choices, negotiating alternatives and changes to commitment, goal, role distribution, etc. It seems desirable to answer this type of questions from a general stand-point and there are examples that suggest it also is feasible.

REFERENCES

- Barker, R. (1968) *Ecological psychology: Concepts and methods for studying the environment of human behavior*. Stanford CA: Stanford University Press.
- Becchio, C., Bertone, C., & Castiello, U. (2008) How the gaze of others influences object processing. *Trends in Cognitive Science*, 12, 254-258.
- Becchio, C., Sartori, L., & Castiello, U. (2010) Toward you: The social side of actions. *Current Directions in Psychological Science*, 19(3), 183-188.
- Bratman, M.E. (1993). Shared intention. *Ethics*, 104, 97–113.
- Bratman, M.E. (1999). *Faces of intention: Selected essays on intention and agency*. Cambridge: Cambridge University Press.
- Brinck, I. (2014) Developing an understanding of social norms and games: Emotional engagement, nonverbal agreement, and conversation. *Theory and Psychology*, 24(6), 737-754.
- Brinck, I. (2015) Understanding social norms and constitutive rules: Perspectives from developmental psychology and philosophy. *Phenomenology and the Cognitive Sciences*. 14(4), 699-718.
- Brinck, I., Balkenius, C., & Johansson, B. (2016) Making place for social norms in the design of human-robot interaction. In: *What Social Robots Can and Should Do. Proceedings of Robophilosophy/TRANSOR 2016*. Eds. Seibt, J. Nørskov, M., Schack Andersen, S. Vol. 290, Frontiers in Artificial Intelligence and Applications. IOS Press: Amsterdam, pp. 303-312.
- Brinck, I., Reddy, V., & Zahavi, D. (2017) The Primacy of the We? In: *Embodiment, Enaction, and Culture. Investigating the Constitution of the Shared World*. Eds. C. Durt, T. Fuchs, & C. Tewes. MIT Press: Cambridge Mass, pp. 131-147.
- Butterfill, S. A. (2011) Joint action and development. *The Philosophical Quarterly*, Vol.62 (No.246). pp. 23-47
- Carpenter, M., & Liebal, K. (2011). Joint attention, communication, and knowing together in infancy. In A. Seemann (Ed.), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience* (pp. 159–181). Cambridge, MA: MIT Press.
- Dale, R., Fusaroli, R., Duran, N., & Richardson, D. C. (2014) The self-organization of human interaction. *Psychology of Learning and Motivation* 59:43–96.
- Gallagher, S., & Marcel, A.J. (1999). The self in contextualized action. *Journal of Consciousness Studies*, 6, 4–30

- Gallagher, S. (2008) Direct perception in the intersubjective context. *Consciousness and Cognition*, 17 (2):535-543
- Gilbert, M. (1989). *On social facts*. Princeton, NJ: Princeton University Press.
- Heft, H. (2007) The social constitution of perceiver-environment reciprocity. *Ecological Psychology*, 19(2):85- 105.
- Hutchins, E. (1995) *Cognition in the wild*. Cambridge MA: MIT Press.
- Hobson, R.P. (2005) What puts the jointness into joint attention? In: Eilan N, Hoerl C, McCormack T, Roessler J (eds) *Joint attention*. Oxford: Oxford University Press, pp 185-204
- Krueger, J. (2012) Seeing mind in action. *Phenomenology and the Cognitive Sciences*, 11 (2):149–173
- Malafouris, L. (2013) *How things shape the mind. A theory of material engagement*. MIT Press, Cambridge MA
- Schmidt, R.C., & Richardson, M.J. (2008) Dynamics of interpersonal coordination. In: Fuchs A, Jirsa V (eds)
- Michael, J. (2011) Shared emotions and *joint action*. *Review of Philosophy and Psychology*, 2 (2), 355-373
- Michael, J, Sebanz, N., & Knoblich, G. The Sense of Commitment: A Minimal Approach. *Frontiers in Psychology*. 2015;6:1968.
- Pacherie, E. (2012) The Phenomenology of Joint Action: Self-Agency vs. Joint-Agency. In Seemann (ed.), *Joint Attention: New Developments*. Cambridge MA: MIT Press.
- Coordination: neural, behavioral and social dynamics*, Springer-Verlag, Heidelberg Germany, pp 281-308.
- Schilbach, L., Timmermans, B., Reddy, V., Costall, A. Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, 36(4), 393-414.
- Tollefsen, D., & Dale R. (2013) Naturalizing joint action: A process-based approach. *Philosophical Psychology*, 25(3), 385– 407.
- Tomasello, M., & Carpenter, M. (2007). Shared intentionality. *Developmental Science*, 10(1), 121–12.
- Tuomela, R. (2005). We intentions revisited. *Philosophical Studies*, 125(3), 327–369.

Money as an Interactive and Subjective Social Kind

EYJA BRYNJARSDÓTTIR

University of Iceland

ABSTRACT

In this paper I consider money as a socially constructed kind. I consider the matter of whether and in what sense being socially constructed makes money unreal or “less real” than things that are not socially constructed. Money clearly affects people’s lives in countless ways, people strive to earn it and those without it suffer. So how could money be less than real? Of course it matters what is meant by the term ‘real’ in this context; what people often seem to mean when they make a claim to the effect that money is not real is that it is subjective or mind-dependent. While some have assumed that things that are socially constructed must be subjective¹, others have argued that at least many social constructions can be very real and objective.² This matter is far from straight-forward when it comes to social phenomena such as money. It complicates matters even further that the terms ‘subjective’ and ‘objective’ have many different meanings and it may even be unclear to those applying them at a given time how to best keep track of their intended use of the term.

Different views have been put forth on the subjectivity of social kinds. John Searle claims, for example, that social kinds are ontologically subjective yet epistemically objective. By this he means that the existence of, say, money depends on our mental attitudes (making it ontologically subjective) while the truth of judgments about whether particular pieces of paper qualify as money is independent of our attitudes (making it epistemologically objective).³ An example of a different approach is Muhammad Ali Khalidi’s, who claims that social kinds may very well be ontologically objective, as not all types of mind-dependence yield subjectivity.⁴

While I agree with Khalidi’s take on what kind of mind-dependence is required for ontological subjectivity, I argue that we have strong reasons to consider money an example of an ontologically subjective kind. The basis for my argument is to be found in the distinction between so-called interactive and indifferent kinds, famously made by Ian Hacking. Interactive social kinds are those where the individuals are influenced by the very classification and in return influence and reinforce the classification, such that a feedback loop is formed, or what has been called a looping effect. Hacking claimed that only human kinds were interactive and other social kinds were thus ‘indifferent’.⁵ Others, such as Khalidi, have argued that the looping

1 Ted Sider (2011), *Writing the Book of the World*, Oxford: Oxford University Press.

2 Sally Haslanger (2011), “Ideology, Generics, and Common Ground”, Charlotte Witt (ed.), *Feminist Metaphysics*, Springer Verlag, 179-207; Esa Diaz-Leon (2013), “What is Social Construction?”, *European Journal of Philosophy* 21(2).

3 John Searle (1995), *The Construction of Social Reality*, New York: Free Press.

4 Muhammad Ali Khalidi (2013), *Natural Categories and Human Kinds. Classification in the Natural and Social Sciences*, Cambridge: Cambridge University Press.

5 Ian Hacking (1999), *The Social Construction of What?* Cambridge, MA: Harvard University Press.

effect can take place in various other social kinds that are non-human.

I argue that the looping effect is in place in the case of money, making it an interactive social kind. My argument for this is built upon two main issues: 1) While money itself is of course not a sentient being whose behavior can be influenced by a classification or other social attitudes toward it, it is used and handled by such beings as are all the rules about how money is used. The behavior of human users of money, or what we might call financial behavior, takes place in a context of social power relations as well as rules and regulations about money. At the same time the rules and regulations constitutive of money, as well as the power relations affecting and affected by financial behavior may change because of various aspects of financial behavior. 2) Important analogies can be drawn in the relevant respects between money and human categories where the looping effect is obviously in effect, such as gender categories. While the human categories are obviously quite different from money, I argue that the elements that gives them the looping effect are also in place in money and financial behavior.

Finally, I argue that interactive social kinds are ontologically subjective.

The status function account of human rights

– Too institutional, or not institutional enough?

ÅSA BURMAN

Stockholm University

ABSTRACT

Social ontologists have shown an increasing interest in moral phenomena like human rights. But, can theories developed to explain the nature and structure of the social world explain moral phenomena? One specific example of this general question is the puzzle:

1. The universal right to free speech did not exist before the European Enlightenment, at which time it came into existence.
2. The universal right to free speech has always existed, but this right was recognized only at the time of the European Enlightenment.⁶

This puzzle draws on two common and conflicting intuitions about human rights: The human right to free speech can exist independently of institutions—these institutions simply recognizes rights we already have. In contrast, this right exists because institutions, or the law, says so. Natural rights theorists solve the puzzle by rejecting (1) and accepting (2), while institutional accounts accept (1) and reject (2). One social ontologist, however, wants to show that we can have it all and accept both (1) and (2). John Searle argues that his status function account of human rights can preserve both intuitions by showing that the inconsistency between (1) and (2) is merely apparent. The status function account of human rights is the view that human rights are deontic powers deriving from status functions, just like other institutional rights—rights of private property and citizenship. The challenge for this institutional account is thus to make sense of (2).

I criticize the status function account of human rights by showing why Searle's solution to the puzzle is inadequate. His solution consists in showing that both human rights and other status functions can exist without collective recognition. But, as I will argue, this solution works only for specific instances of human rights (tokens) but not for human rights as a kind (type). Yet, the conflict and puzzle concerns human rights as a kind. There is, then, still an asymmetry between human rights and other status functions regarding types (if we wish to say that human rights as a kind can exist without collective recognition since kinds of status functions cannot). And so, the status function account fails to make sense of our conflicting intuitions about human rights. In short, I will show that the third solution to the puzzle – accept (1) and (2) – is unviable. Are there other ways to solve the puzzle within this framework?

All critics, and Searle, accept (2). But maybe the account is not institutional enough? Another option, given an institutional account, is to deny (2). Susan James, for example, does so in arguing for viewing human rights as “effectively enforceable claims”. This opens up for the objection that it is too institutional since many people do not have the human right to free speech (if it is not effectively enforceable on James's view, or if we change the constitutive rules

6 John Searle, *Making the Social World*, 2010, p. 177.

on Searle's view). Still, this does not preclude us from saying that it is morally wrong to deny people the right to free speech (for other reasons). But it cannot be a rights violation (since the right does not exist in the first place if it is not effectively enforceable or there is no such constitutive rule in place). Despite this problem, I think it is most plausible to view the status function account as a new contender within this institutional camp.

And for those of us who still wish to accept (2), I propose another way to resolve the conflict: The puzzle is about two different senses of a "right": (1) refers to the institutional right; and (2) refers to the moral right. The latter is probably best left to moral philosophers, while the status function account can clarify the nature and existence of the former. This is not a small thing, especially not if human rights today "are the rights of the lawyers, not the philosophers".⁷

REFERENCES

- Feinberg, Joel, "The nature and value of rights" *Journal of Value Inquiry* 4 (4): 243-260 (1970).
- Fotion, Nicholas, "Review: Searle on human rights", *Analysis*, Vol 71 (4) (2011)
- Hart, H.L.A., "Are there any natural rights?" *The Philosophical Review*, 1955 (64:2), 175-191.
- Hindriks, Frank, "Restructuring Searle's Making the social world", *Philosophy of the Social Sciences* 43 (3): 373-389 (2013).
- James, Susan, "Rights as enforceable claims". *Proceedings of the Aristotelian Society*, New Series, Vol. 103 (2003)
- Lobo, Gregory, "A critique of Hindriks' Restructuring Searle's Making the social world", *Philosophy of the Social Sciences*, 2015, Vol. 45(3): 356–362.
- Nickel, James, W., *Making Sense of Human Rights*, 2007.
- Searle, John, *Making the Social World: The Structure of Human Civilization* (Oxford: Oxford University Press, 2010).
- Searle, John, *The Construction of Social Reality*, New York: Free Press 1995.
- Thomasson, Amie, "Foundations for a Social Ontology", *ProtoSociology*, 18-19 (2003), 269-290.
- Tuomela, Raimo "Searle's new construction of social reality" *Analysis*, Vol 71(4), (2011).

7 James W. Nickel, *Making Sense of Human Rights*, 2007.

Normative Representations and Emotions

ANTONELLA CARASSA⁸, MARCO COLOMBETTI^{9,2}

ABSTRACT

A distinctive feature of human beings is their ability to enter normative relations (like obligations, claims, and the like) with their conspecifics. Presumably, the function of normativity is to support human cooperation, broadly construed to include (besides cooperation properly so called) the creation and maintenance of interpersonal relationships and complex social structures. This brings to the foreground the problem of understanding what human mental capacities make normativity possible. Many contributions dealing with such capacities (mainly in the moral domain) are available in the literature, one notable example being Tomasello's recent work on the psychological infrastructure of morality (Tomasello, 2016). Remarkably, Tomasello's account almost completely neglects the affective infrastructure, and in particular moral emotions, which on the contrary are considered crucial by many scholars of moral psychology (e.g., Haidt, 2003). In our presentation we shall argue that understanding the relationships between cognitive and affective aspects is critical to the development of a satisfactory theory of normativity.

We contend that normative relations between agents involve mental representations of a special kind. Differently from the most widely analysed "first-personal" representations (like individual beliefs, desires, intentions, and so on), normative representations are intrinsically *second-personal*, in the sense of making an agent responsible or accountable to another agent for doing or not doing something (Darwall, 2006; Carassa & Colombetti, 2014): every obligation is an obligation to someone, every right is a right against someone, and so on. Like desires and intentions, normative representations have a world-to-mind direction of fit; but differently from desires and intentions, they make an agent accountable to a *second party* for their satisfaction. Every second-personal relation can be viewed from the perspective of each of the two parties. In the case of normative relations, this gives rise to the fact, already remarked by Hohfeld (1923) in the legal domain, that to a normative relation binding *A* to *B* there corresponds a correlative normative relation binding *B* to *A*. This implies that the mental representations of a party of a normative relation intrinsically represents also the correlative relation. For example, if Ann regards herself as obligated to water Bob's flowers while he is away, she *ipso facto* represents Bob as having claim against her that she so does. This does not mean that Bob actually holds the matching mental representation: at least in principle, an agent may hold a normative representation binding her to another agent independently of any normative representation held by the latter and binding him to the former. However, the representations held by the parties of a normative relation are functionally interdependent in that, at least in general, the functional goal of normativity is best served if the representations held by both parties match: this means, for example, that if Ann represents herself as obligated to Bob to water his flowers, then Bob's represents himself as having claim against Ann that she so does (and vice-versa).

8 Faculty of Communication Sciences, Università della Svizzera italiana, Lugano, Switzerland {antonella.carassa,marco.colombetti}@usi.ch

9 Department of Electronics, Information, and Bioengineering, Politecnico di Milano, Milano, Italy marco.colombetti@polimi.it

To serve their function, normative representations need to motivate humans to interact in the right ways. Certain emotions (like gratitude, resentment, guilt, and the like) are known to contribute significantly to achieving this result by: (i), motivating the subject of the emotion to behave in certain ways; and (ii), by eliciting reactive emotions in other agents, which in turn will be affected in their behaviour. But this brings in an interesting issue. On the one hand, normative representations have a world-to-mind direction of fit, and thus fulfil their function by contributing to bringing about their conditions of satisfaction. On the other hand, certain types of emotions (which we call *normative emotions*) are known to affect the behaviour of agents bound by normative relations. We can therefore expect that systematic relationships hold between normative representations and the emotions that these representations elicit in specific situations.

Normative emotions are subjective experiences that assess specific situations as *compliant* or *noncompliant* with certain normative representations: while “positive” emotions of righteousness, gratitude, etc., support the production of compliant behaviour, “negative” emotions of resentment, indignation, guilt, shame, etc., contribute to the management of violations. From this point of view normative emotions may look similar to basic emotions (like anger, disgust, fear, etc.), which assess the valence of a state of affairs with respect to some fundamental biological concern. Contrary to basic emotions, however, normative emotions have the distinctive feature of assessing a state of affairs *relative to a second party*: one feels guilty to someone about something, has resentment *towards someone* for something, and so on. This property of normative emotions reflects the fact that normative relations are intrinsically second-personal.

A further property of normative emotions is that they involve physical or behavioural changes that can be perceived by other agents. This is often the case also for those emotions, like fear, whose function is not primarily social; however, in the case of normative emotions this feature is crucial, because the function of such emotions is to shape the interactions of agents bound by normative relations that make them accountable to each other. For example, suppose that Ann did not water Bob’s flowers, in the context of an obligation to Bob to do so. Due to the second-personal nature of normative relations, recognising that something went wrong is not a private issue of Ann and Bob separately from each other (as it would be, for example, if watering Bob’s flowers were a personal goal of Ann’s and, separately, a private desire of Bob’s). On the contrary, it is part of the very idea of accountability that Ann and Bob make their own attitudes towards the situation accessible to each other: by displaying some degree of guilt, Ann shows to Bob that she still recognises her obligation; symmetrically, by displaying some degree of resentment, Bob shows to Ann that he took her obligation seriously.

REFERENCES

- Carassa, A., & M. Colombetti (2014). Interpersonal responsibilities and communicative intentions. *Phenomenology and the Cognitive Sciences*, 13, 145–159.
- Darwall, S. (2006). *The second-person standpoint: Morality, respect, and accountability*. Harvard University Press
- Haidt, J. (2003). The moral emotions. In R.J. Davidson, K.R. Scherer, & H.H. Goldsmith, eds., *Handbook of affective sciences*, 852–870. Oxford University Press.
- Hohfeld, W.N. (1923). *Fundamental legal conceptions as applied in judicial reasoning*. Cook.
- Tomasello, M. (2016). *A natural history of human morality*. Harvard University Press.

Ghost Stories. Toward an Hegelian Approach to Social Ontology

ALESSANDRO DE CESARIS

University of Eastern Piedmont (Vercelli) and Humboldt Universität zu Berlin

ABSTRACT

In a famous paper Andy Clark and David Chalmers introduced the idea of an extended mind in order to explain the nature of some technological objects. According to their position, that they define as a form of “active externalism”, human mind is neither confined inside our skull nor limited by our skin, and the process of cognition can’t be reduced to conscience, but it can rather be transferred to other devices (as computers, phones, neural implants). A relevant aspect of this theory and that it entails a deep redefinition of a vast variety of objects from the ontological standpoint: in order to grasp this theory we must in fact not only modify our naive conceptions about the nature of mind, but also and most importantly reshape our conception of many technological instruments that are directly involved in our active cognitive processes. Clark and Chalmers’ proposal doesn’t add much to the results already achieved by philosophical anthropology and media theory during the 20th century, and yet it has introduced the question concerning technological objects and prostheses in the area of philosophy of mind. Some scholars have remarked that the theory of extended mind can’t be limited to technological objects such as phones, computers and similar devices: streets, billboards, buildings, infrastructures, texts, contracts, any form of technical object is as a matter of fact an extension of the kind addressed by Clark and Chalmers. From this standpoint, the theory of extended mind can be seen as the starting point for a general social ontology. It has already been remarked that the notion of “extended mind” resembles quite closely the Hegelian notion of “objective spirit”, and that it lets us extend this very notion outside of the boundaries set by Hegel himself (who considered just social phenomena as families, civil society, corporations, markets, contracts, states and so on). And yet, the notion of “objective spirit” (objektiver Geist) seems to be somehow obsolete. This impression is owed to a certain spiritualistic and metaphysical understanding of Hegel’s thought. Rather than that, the notion of objective spirit allows a logical-ontological foundation of the question concerning media in its most general and comprehensive form. Today philosophy of media isn’t yet an autonomous and well-defined discipline, and in this landscape Hegel’s thought can grant a theoretical foundation to it, since it offers a social ontology that is structurally based on a general ontology and on a logic, and at the same time allows to consider the question concerning media on all its different levels: metaphysical, anthropological, social, aesthetic.

Various philosophers have already remarked the interest of the category of objective spirit for social ontology. In my paper I do not intend to simply mention this concept as an interesting curiosity about the prehistory of social ontology, but rather I’d like to ask about the condition of possibility, the methods and the potentialities of an Hegelian approach to social ontology itself. The main focus of such an approach would not be, as traditionally believed, the endorsement of a “dialectical” logic instead of classic formal logic, but rather the discussion and the extension

of the concept of “objective spirit”, notion that has to be freed from any spiritualistic and metaphysical misunderstanding.

I intend to argue that Hegel’s philosophy of objective spirit is a social ontology that can immediately be interpreted as a media theory. Considering Marshall McLuhan’s famous definition of media as extensions of human body, Hegel’s notion of “objective spirit” reveals the insufficiency of Clark and Chalmers’ theory: objective spirit is in fact at the same time extended mind and extended body, and the very notion of “spirit” questions the possibility to hold an ontological distinction between mind and body with respect to social objects. Hegel’s philosophy of objective spirit doesn’t limit its account of social objects to cognitive processes, but rather contemplates every social dimension, from contracts to technological objects, from infrastructures to social groups and institutions. On a more general note, a social ontology based on such an approach would be naturally and immediately connected with a general ontology, a theory of art and anthropology without losing its own autonomy.

The choice to understand Hegel’s social ontology as a media theory seems to be particularly justified if one considers that the notion of “medium” in itself rejects the classical mind/body opposition. From this standpoint, aim of my paper is to show that Hegel’s concept of “spirit” is part of a tradition of concepts, whose function is exactly to avoid this opposition and to address a very peculiar dimension, that of social phenomena. For this reason I intend to analyze the concept of spirit on the basis of Giorgio Agamben’s analysis of the notion of “phantasma” in Western thought. This analysis has been recently extended by Fabián Ludueña Romandini and turned into a proper social ontology in the form of a spectrology. The classical notion of phantasma, already mentioned by some theorists (like Pierre Levy) as a core mediological concept, allows us to understand the peculiar status of what Hegel defines “spiritual”: neither subject nor object, neither abstract idea nor concrete thing, it is the spectre that defines the field of social ontology.

SHORT BIBLIOGRAPHY

- Agamben G., *Stanzas: Word and Phantasm in Western Culture*, translated by W.I. Martinez, University of Minnesota Press, 1992.
- Clark A., Chalmers D., *The extended mind*, “Analysis”, 58, 1, 1998, pp. 7-19.
- Crisafi A., Gallagher S., *Hegel and the extended mind*, “AI&Society”, 25, 1, 2010, pp. 123-129.
- Ferraris M., *Documentality. Why It Is Necessary to leave traces*, translated by R. Davies, Fordham University Press, 2012.
- Hegel G.W.F., *Encyclopedia of the Philosophical Sciences. The Philosophy of Mind*, translated by W. Wallace and A.V. Miller, Clarendon Press, 1971.
- Hegel G.W.F., *Outlines of the Philosophy of Right*, translated by S. Houlgate, Oxford University Press, 2008.
- Levy P., *Collective intelligence*, translated by R. Bononno, Basic Books, 1999.
- Romandini F.L., *Principios de Espectrología. La comunidad de los espectros II*, Mino Dávila, 2016.

Taking one for the team. Team-reasoning and the possibility of collective alief

BO ALLESØE CHRISTENSEN

Aalborg University

ABSTRACT

This presentation will focus on the structure of belief-alief argued for by Gendler (2008), asking how this could inform a sense of team reasoning. Three examples will set the discussion. Imagine a combat situation in the first world war. War bunkers set on either side of the war zone, people firing guns and throwing hand grenades at each other. Now one such hand grenade fulfils its target and touches ground in one of the bunkers, with the result that one of the persons in the bunker throws himself on top of the grenade to protect the rest of people in the bunker. Is this in congruence with the normal way of thinking of team-reasoning? The guy throwing himself on the grenade most likely believes he ought to do what he is about to do, but he is probably also alieving something carrying the content like “Danger. This will kill me. Gotta get away.” Another example, imagine a group of kids playing football on a street. By accident one of the kids kicks the ball so hard, it breaks a window. Obviously, all of the kids are somehow responsible by partaking in the football game, but only the kid who kicked the ball, or in case this kid has been breaking too many windows, perhaps one of the other kids will be claiming responsibility. Approaching the house with the broken window the boy believes he is taking one for the team, but most likely he is wishing he was somewhere else. The last example is one Gendler presents, but we will use it here with a slight alteration. In 2007, a tourist attraction on the western rim of the Grand Canyons, known as Grand Canyon Skywalk, opened. It consists of a horseshoe-shaped cantilevered glass walkway placed 4000 feet above the Grand Canyon, and extending 70 feet from the Canyon’s rim. Gendlers point is, that visitors to the skywalk surely believe that the walkway will hold when stepping on to it. But they also alieve something very different: “The alief has roughly the following content: Really high up, long way down. Not a safe place to be! Get off!!” (Gendler 2008, 635).

Now Gendler presents this primarily from a first-person perspective, but we can easily imagine two people taking a stroll on the Skywalk, with only one of them really wanting to. The other tags along but as gesture of upholding a sense of “we”: could be out of love for the other, a grandparent protecting their son or daughter’s offspring hence the family, or even out of duty for someone working in eldercare, because “this is what we activity assistants do”. Each one is “taking one for the team” believing this is the right thing to do, while experiencing what Gendler (2008, 641) terms a belief-discordant alief, with a representational-affective-behavioural content: the visual appearance of the depth of the Canyon, a feeling of discomfort bordering on fear, and a physical disposition of wanting to retreat. Now two issues interest me with the above examples, and which I will elaborate on. First, I think, the examples make sense as illustrations of what team-reasoning also is about compared to an over-rational or -intellectualised game-theoretical or equilibrium setting ala Sugden (2007). Especially if we, as Sugden, understand this setting as prescribing acting as a team member as

always conditional on the assurance that others have endorsed the same principle. This fails to consider the possibility of agents choosing from a genuinely joint perspective (Schweikard and Schmid 2013), i.e. they do not consider what is best for them individually, given the expected choice of the other. The man jumping on the grenade is not consciously expecting all other people to do the same if they were in his situation. If he did, he most likely wouldn't jump himself. Aliefs are part of the decision-making in the cases above, and help us understand how complex the structure of beliefs involved reasoning and decision making are. Especially they point a way between a too rationally conceived belief structure like Sugdens and conceiving aliefs as irrational (see Nagel 2012). They allow instead the consideration of arational and non-conditional (as not depending on expecting others to do the same) elements as part of the the team-reasoning. Second, while Gendler only considers the belief/alief structure from a first person perspective, the examples above connects this with a joint perspective as well. But how are we to understand this structure from a joint perspective?

Here I will critically pursue two directions in Gendlers argument which could open up for considering a we-persepective. First, Gendler claims that aliefs can drive action independently of desire. This puts alief in contrast to belief which ordinarily is thought to motivate only in conjunction with desire, but it also makes it possible to characterise alief as instigating action from a we-perspective, or, at a minimum, from a perspective not involving individual desires. Second, Gendler thinks that the representational component in aliefs are not subject to accuracy in the sense in which beliefs are, namely able to distinguish between appearance and reality. Both these conditions would have to be included when describing taking one for the team as an example of collective alief.

REFERENCES

- Gendler, T. S. (2008) Alief and Belief. *The Journal of Philosophy*, 105(10): 634-663
- Nagel, J. (2012) Gendler on Alief. *Analysis* 72(4): 774-788
- Schweikard, David P. and Schmid, Hans Bernhard (2013) "Collective Intentionality", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/sum2013/entries/collective-intentionality/>>.
- Sugden, R. (2007) The logic of team-reasoning. *Philosophical Explorations* 6(3): 165-181

When the (Democratic) State Acts, does its Citizenry?

STEPHANIE COLLINS AND HOLLY LAWFORD-SMITH

University of Manchester / University of Melbourne

ABSTRACT

In his first week as President of the United States, Donald Trump made numerous executive orders, one of which was a 120-day suspension of the US's refugee programme and a 90-day ban on travel to the US for citizens of seven countries. This order led to various travellers (including visa holders) being detained, denied entry, and refused travel. It deeply affected those refugees who otherwise would have been admitted to the US, at least some of whom might now not be able to travel there at all. In short, the order caused harm. Who did the harm? In this paper, we examine one possible answer: 'the enfranchised citizenry.' We are interested in the following:

Target Claim: the enfranchised citizenry taken as a whole, and each enfranchised citizen, acts when the state does.

There are several literatures that bear on this claim, but none directly give a verdict on its truth or falsehood. For example, theories of collective agency do not settle this directly. Early theorists of collective agency took small, usually more-or-less egalitarian, groups of people as their targets. Some authors have been explicit about whether they think their accounts 'scale up' to larger, and not necessarily egalitarian, groups (e.g., Margaret Gilbert). Others have explicitly bracketed such groups (e.g., Michael Bratman). More recent theories (e.g., Christian List and Philip Pettit's) have been designed to apply to large-scale and hierarchical groups such as states, but these theories do not provide clear guidance on what it takes for an individual to count as part of such a group agent.

That said, theories of collective agency articulate the sense in which the state is an agent: the state is an agent by virtue of possessing its own rational decision-making procedure, which it uses to generate its distinctive rational point of view—its own distinct bundle of beliefs and desires, from which it acts and which it seeks to maintain. Such agents have control over their actions, in the sense that if the agent uses its decision-making procedure to decide to do something, then that thing usually gets done because of the agent's decision. They tend to face a wide range of options for deliberation. Group agents can also put options before themselves, that is, the group can decide which options, and how many, its decision-making procedure will deliberate over in the first place. And they can do all this robustly; that is, these abilities are not hostage to minor changes in their circumstances. Also, group agents are a source of unity for their members: they provide overarching goals, procedures, and structures that bind members together. Yet their decisions (and subsequent actions) supervene on various facts about members, which implies that facts about members influence facts about the group agent. Yet merely having influence is not sufficient for being a member—after all, the weather and the laws of physics also influence facts about group agents. Members seem to be distinguished from the weather by the fact that their influence is mandated by the collective—it

occurs ‘from the inside’ of the group agent—and is in some sense voluntary or willed.

This brief characterisation raises at least seven features of group agents that bear upon our target claim. Each feature is scalar: it can be satisfied to a greater or lesser extent. If a democratic citizenry exhibits a higher number of these features, and to a greater extent, then the case for the target claim is stronger, *vis-à-vis* that citizenry. These features are: control, scope, directness, robustness, unity, influence, and voluntariness. That is to say, the case for a given enfranchised citizenry constituting its state is strongest when that citizenry has control over what the state does, when that control can be exercised over a sufficient range of options, when no factors intervene between the citizenry’s decision and its enactment, when the citizenry’s control is not precarious, when the state unifies its citizens under certain goals or procedures, when each citizen can influence or make a difference to what the state decides, and when each citizen’s use of their mandated influence is willed or voluntary.

In this paper, we group these features into four desiderata: joint control (including the particularly, scope, directness, and robustness of that control), unity, influence (including difference-making and equality of influence), and voluntariness. The first two desiderata—control and unity—bear upon the question of whether the enfranchised citizenry, taken as a whole, acts when the state (considered as an agent) acts. The latter two desiderata—influence and voluntariness—bear upon whether any given individual citizen acts when the state (considered as an agent) acts.

There is much to be said, and no easy conclusion. We plan to give the fullest possible articulation of these desiderata, explaining how each can be used to both support, and oppose, the target claim, and leave the reader to decide which bullets they’re happy to bite. We work through each of the four desiderata in turn, motivating their importance for the target claim and explaining how they work in favour of or against the target claim. One of the authors is convinced that the target claim is false, and that we must look elsewhere than the citizenry for an account of who or what acts when the state does; the other author is convinced that the target claim is true, so there’s no need to look elsewhere. We hope to make use of our disagreement in illuminating the target claim. On many of our desiderata, it’s likely that some democratic citizenries will satisfy the desiderata while others will not. Instead of making definite judgments on particular citizenries (such as the United States), our aim is to motivate and explain the desiderata and discuss the extent to which they are satisfied by citizenries of liberal democracies in general.

Collective reasoning and collective responsibility gaps

HEIN DUIJF
Utrecht University

ABSTRACT

How should we reason in interactive decision contexts? Can we specify the difference between non-cooperative and cooperative decision problems? I aim to shed some light these questions by contrasting several practical reasoning paradigms. Bacharach (2006, Ch. 1) and Sugden (2000, Sections 2, 3, 7 and 8) argue that traditional game and decision theory needs to be augmented with a collectivistic reasoning method to successfully address cooperation problems like the Hi-Lo game. They introduce the team reasoning account of cooperation, which appeals to the *reasoning method* by which an individual agent reasons about what to do. An individual agent engaged in team reasoning “works out the best feasible combination of actions for all the members of her team, then does her part in it” (Bacharach, 2006, p. 121). The team reasoning literature has adopted the reasoning schema in Figure 1 (where (1)–(6) denote the premises, and • states the conclusion).

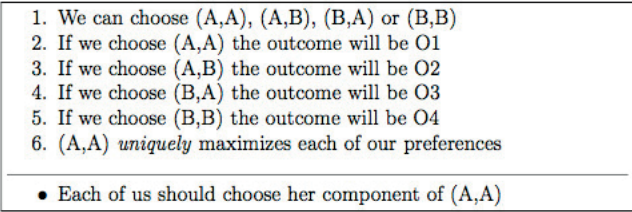


Figure 1: Team reasoning schema.

		P2	
		high	low
P1	X	(2,2)	(0,0)
	Y	(2,2)	(0,0)
	low	(0,0)	(1,1)

Figure 2: The alternative Hi-Lo game.

This team reasoning schema has a problem: it fails to deliver the desired recommendations in the *alternative* Hi-Lo game depicted in Figure 2. It seems that a good theory of cooperation should recommend choosing *high* to P2 and should recommend P1 to choose either *X* or *Y*, rather than *low*. Team reasoning, unfortunately, fails to deliver any of these recommendations.

The problem for the team reasoning schema is that premise (6) is undermined: there are multiple group actions that maximize each of our preferences.

Our diagnosis of the failure of team reasoning is twofold: First, premise (6) in the team reasoning schema (Figure 1) highlights that it relies on an unrealistic uniqueness assumption, namely that there is a *unique* group action that maximizes each of the members' preferences. To address the alternative Hi-Lo game, we need to drop this problematic assumption. Second, the team reasoning schema blurs the relation between team reasoning on the one hand, and individual and collective reasoning on the other. I aim to contribute to theories of cooperation by augmenting the team reasoning paradigm to, what I call, participatory reasoning in order to address these problems.

The participatory reasoning schema depicted in Figure 3 marks my main contribution (where (I●) denotes the conclusion; (C●) states intermediate conclusions; and the others denote the premises).

It is important to note that the participatory reasoning schema delivers the desired recommendations in the alternative Hi-Lo game (Figure 2). If P2 adopted the participatory reasoning schema, her reasoning would go as follows: first she reasons at the collective level and concludes that we should choose $(X, high)$ or $(Y, high)$. Then she reasons at the individual level and concludes that she should choose *high*, because choosing *high* is compatible with a best group action whereas choosing low is not. Similarly, participatory reasoning recommends P1 to choose either X or Y. My participatory reasoning schema thus resolves this paradox of team reasoning.

<p><i>(Stage 1: the collective level)</i></p> <p>(C1) We can choose (A,A), (A,B), (B,A) or (B,B)</p> <p>(C2) If we choose (A,A) the outcome will be O1</p> <p>(C3) If we choose (A,B) the outcome will be O2</p> <p>(C4) If we choose (B,A) the outcome will be O3</p> <p>(C5) If we choose (B,B) the outcome will be O4</p> <p>(C6) We prefer O1 over O2, O3 and O4</p> <hr/> <p>(C●) We should choose (A,A)</p>
<p><i>(Stage 2: the individual level)</i></p> <p>(I1) I can choose A or B</p> <p>(I2) If I choose A the outcome will be that we choose (A,A) or that we choose (A,B)</p> <p>(I3) If I choose B the outcome will be that we choose (B,A) or that we choose (B,B)</p> <p>(C●) We should choose (A,A)</p> <hr/> <p>(I●) I should choose A</p>

Figure 3: Participatory reasoning schema.

Can this mode of practical reasoning shed light on responsibility gaps in collective action contexts? In case of a collective wrongdoing this triggers the study of which (if any) members can be held responsible for it.¹⁰ I therefore focus on cases of collective failure, which, in our current framework, are cases where an inferior group action is performed. If we think of responsibility as performing those actions that are admitted by the participatory reasoning schema, then we thus investigate the ways in which individuals that endorse the participatory reasoning schema to maximize their collective preference can still end up performing an inferior group action.

In such a case, my set-up provides a distinction between two possible diagnoses: (1) The collective wrongdoing could originate from diverging expectations regarding *out-group* members. In this case, the group members with inaccurate expectations seem to be at fault. There is no responsibility gap.¹¹ (2) The inferior group action may result from diverging expectations regarding *in-group* members. In this case, it could be impossible to say which, if any, group member is at fault. There is a responsibility gap: there is a collective wrongdoing although none of the members can be held responsible.

REFERENCES

- Michael Bacharach. 2006. *Beyond Individual Choice: Teams and Frames in Game Theory*. Princeton University Press, Princeton.
- Matthew Braham and Martin van Hees. 2011. Responsibility voids. *The Philosophical Quarterly* 61, 242 (2011), 6–15.
- Tracy Isaacs. 2011. *Moral Responsibility in Collective Contexts*. Oxford University Press.
- Christopher Kutz. 2000. *Complicity: Ethics and Law for a Collective Age*. Cambridge University Press, Cambridge.
- Robert Sugden. 2000. Team preferences. *Economics & Philosophy* 16, 2 (2000), 175–204.

¹⁰ It is disputed whether this is possible, compare Kutz (2000) and Isaacs (2011).

¹¹ Braham and van Hees (2011, p. 13) provide a more nuanced assessment: “we may have to say that *m2* and *m3* are not responsible for the state of affairs as such – and hence that there is still an epistemic void – although both are blameworthy for holding unjustified beliefs.”

Causal Responsibility for Emissions in Global Supply Chains

ANTON ERIKSSON

University of Sheffield

ABSTRACT

The production and consumption of internationally traded goods plays a large role in the occurrence of climate change. The greenhouse gases that result from manufacturing certain products make up for a considerable amount of total global emissions. Such emissions, in turn, cause harm to current and future generations. This raises the question of where responsibility for this harm should fall. In the debate on moral responsibility for climate change, however, focus has mainly been on discrete sets of agents, like states, companies and individual consumers. This is despite the fact that our world is an essentially interrelated one. An increasing number of products today are manufactured in more than a single country and an extraordinary number of persons are involved in the supply chain of any one product. In determining the causes of climatic harms, it therefore seems crucial to look at how the actions of actors interplay at different stages of production and trade. In this paper, I will look at the role played by those involved in global supply chains in causing harm through emissions. The aim is to establish how causal responsibility for emissions ought to be shared between different actors in the production and consumption of emission-heavy products. I will first look at the relevant type of agents that are active along global supply chains. Second, I will apply a counterfactual theory of causation to this issue and respond to a number of worries related to the problems of redundant causation and joint causation, respectively. Finally, I will briefly discuss the normative implications of this analysis.

Social Cognition, Empathy, and Agent-Specificities in Cooperation

ANIKA FIEBICH

University of Milan

ABSTRACT

In the contemporary debate, joint actions are often defined as any kinds of social interactions in which two (or more) agents share an intention to cooperate; for example, playing in an orchestra, or moving a piano together. It has been discussed controversially what a 'shared intention' or 'we-intention' comes to in the contemporary debate. Searle (1990), for example, has argued for a 'we-intention' as a specific kind of attitude involving collective intentionality in the mind of a single individual; accordingly, an individual can we-intend even if no other agent exists, and 'we-intentions' may occur in a brain in a vat. Bratman (1993), in contrast, has argued for an additive account that defines a 'shared intention' of a group as being composed of the personal intentions of the group members that are interrelated in a specific way. Gilbert (2009) also accounts for a 'shared intention' as being composed of personal intentions, yet as not reducible to them; in fact, a shared intention of a group as a plural subject may continue to exist even if some personal intentions of single group members have been decayed or rescinded. Of course, this list is not exhausted but what these and other accounts in the debate share is a detailed discussion of the nature of 'we-intentions' or 'shared intentions' as well as the interrelation between the shared intention of the group and the personal intentions of the single group members. These approaches fail, however, to account for the social cognitive and affective skills and competencies that are required to be engaged in joint actions involving shared intentions as well as the dedicated role that agent-specificities may play.

But other philosophers and cognitive scientists from other disciplines have discussed the social cognitive prerequisites of Searle's 'we-intentions' or Bratman's 'shared intentions', including the capacity of building meta-representations or shared task representations (Dominey and Warneken 2011) or having a 'theory of mind', i.e. the capacity to understand another person's behaviour by attributing belief-desire pairs to him or her (Tollefsen 2005). Building meta-representations and understanding people's behaviour in terms of beliefs and desires are cognitively rather sophisticated capacities that emerge not until age 4 in human ontogeny (Wellman et al. 2001). Developmental studies have shown, however, that already 2-year-old children are engaged in cooperative problem solving activities and social games involving joint attention (e.g., Warneken et al. 2006).

In my talk, I argue for cooperation as a three-dimensional phenomenon lying on the continua of (i) a behavioural, (ii) a cognitive, and (iii) an affective axis. The developmental findings suggest that an adequate approach to cooperation cannot be reduced to cooperative activities involving shared intentions. Rather, the social cognitive demands involved in cooperative activities are a matter of degree, ranging from cognitively demanding cooperative activities involving shared intentions that presuppose sophisticated social cognitive skills such as having a theory of mind to basic joint actions like intentional joint attention (Fiebich and Gallagher 2013). Hence it is

useful to devote the term ‘joint action’ to cognitively demanding cooperative activities that involve shared intentions, and to use the term ‘joint activities’ for any other cognitively less demanding activities in which agents cooperate with each other to achieve a particular outcome. That is, any cooperative phenomenon can be located on a cognitive axis where the highest point of cognitively demandingness is joint actions involving shared intentions. Moreover, any cooperative phenomenon can be located on a behavioural axis, ranging from highly complex coordinated behaviours (potentially being determined by sophisticated rules and roles) to basic coordinated behaviours such as simple turn-taking activities.

Finally, cooperative activities can be located on the continuum of an affective axis that is determined by the degree of ‘sharedness’ of the affective state in question, ranging from individual empathetic feelings of a single agent to more or less complex shared emotions of agents. Findings from social psychology suggest that the (shared) affective states of the cooperating agents are reciprocally interrelated to the cognitive states of the agents and the coordinated behaviours of the agents. For example, van den Hooff et al. (2012) found that empathy affects eagerness and willingness to share knowledge, and that these emotions, in turn, influence the knowledge of having a shared intention. Finally, findings from social psychology indicate that empathy may play an important role in cooperative activities and that agent-specificities may shape the empathetic experiences of the cooperating agents; e.g., Ford and Aberdein (2015) investigated the influences of empathy on the joint Simon task and found that the magnitude of the joint Simon effect was correlated positively with empathy only when the person who was sitting next to the participant were a friend but not when it was a stranger.

REFERENCES

- Bratman, M.E. (1993). Shared Intention. *Ethics*, 104(1), 97-113.
- Dominey, P., & Warneken, F. (2011). The basis of shared intentions in human and robot cognition, *New Ideas in Psychology*, 29, 260-274.
- Fiebig, A., & Gallagher, S. (2013). Joint attention in joint action. *Philosophical Psychology*, 26(4), 571-587.
- Ford, R. M., & Aberdein, B. (2015) Exploring social influences on the joint Simon task: empathy and friendship. *Frontiers in Psychology*, 6(962),1-11.
- Gilbert, M. (2009). Shared intention and personal intentions, *Philosophical Studies*, 144, 167-187.
- Searle, J. (1990). Collective Intentions and Actions. In Cohen, P., Morgan, J. and Pollack, M. (Eds.), *Intentions in Communication* (pp. 401-415). Cambridge, MA: MIT Press.
- Tollefsen, D. (2005). Let’s Pretend! Joint Action and Young Children, *Philosophy of the Social Sciences*, 35, 75-97.
- Van den Hoof, B., Schouten, A.P., & Simonovski, S. (2012). What one feels and what one knows: the influence of emotions on attitudes and intentions towards knowledge sharing. *JKM*, 16(1), 148 – 158.
- Warneken, F., Chen, F. & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development*, 77, 640-663.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development*, 72, 655-684.

Theories of Group Agency: Preliminary Experimental Data

BRIAN FLANAGAN

Maynooth University

ABSTRACT

It is standardly assumed that a group may act, member disagreement notwithstanding, by forming a will of its own (e.g., Copp 1995; Pettit 2001; Kornhauser and Sager 2004). Likewise, it is standardly assumed that the theory of group agency ought to cohere with our conceptual intuition (e.g., Arrow 1963; Ludwig 2007; List and Pettit 2011). Conversely, there is sharp divergence between most social choice theorists, on the one hand, and most philosophers of action, on the other, over the content of this intuition, and, consequently, over which of two alternative theories ought to be adopted. In principle, appeals to a theory's intuitiveness are amenable to empirical investigation. Nevertheless, the relevant scholarly communities are yet to depart from the traditional approach of 'armchair' reflection supplemented by intra-community dialogue. Two clear avenues for advancing debate therefore present themselves. First, the prospect for a fruitful synthesis of the two leading theories remains to be explored. To this end, I integrate the hitherto insulated literatures on collective intentionality and social choice to characterize a composite theory of group agency. Second, the claimed intuitiveness of these theories might be subjected to experimental confirmation. Having conducted a survey of 180 university students, I report the results of the first systematic assessment.

'Substantive majoritarianism' holds that a group wills a particular act alternative if, and only if, the alternative is majority preferred. The popularity of substantive majoritarianism among social choice theorists is reflected in their common assumption that: [It] surely violates minimal democracy... [if]... there is an alternative that is strictly preferred by two of the three individuals (Austen-Smith 2006, 908; similarly, Riker 1980, 443). In contrast, 'procedural unanimitarianism', holds that a group wills a particular alternative if, and only if, it possesses a unanimously preferred relation to member substantive preferences (e.g., Tuomela 1995; Copp 1995; Bratman (forthcoming)). A synthesis of substantive majoritarianism and procedural unanimitarianism is available, namely, 'procedural majoritarianism', the theory that a group wills a particular alternative if, and only if, it possesses a majority preferred relation to member substantive preferences.

The reported survey is an example of a branch of experimental philosophy that is variously characterized as, 'the verification project' (Pust 2012), and, 'experimental analysis' (Nadelhoffer and Nahmias 2007). As such, it faces four general sorts of methodological challenge, concerning, respectively, the problems of question inclusivity (e.g., Marti 2009), expert-lay disagreement (e.g., Williamson 2011), guesswork answering (e.g., Bengson 2013), and ambiguity (e.g., Deutsch 2009). Having addressed how my survey design and interpretation meets these challenges, I present the results. I find that these undermine the claimed intuitiveness of both leading candidates, namely, substantive majoritarianism and procedural unanimitarianism, whilst they substantiate the intuitiveness of their synthesis, procedural majoritarianism.

Organizations: an Ontological Approach

FRANCESCO FRANDA

University at Buffalo

ABSTRACT

Every day, we deal with organizations of various kinds (such as companies, schools, hospitals, and governments), and these interactions deeply shape our lives. Hence, an important part of social science is devoted to the study of the mechanisms underlying these various types of organizations. However, this focus does not necessarily translate into a precise definition of the domain of inquiry. In my talk, I propose a definition of ‘organization’ and I present an ontology of organizations. In my work, I employ Basic Formal Ontology (BFO), a top-level ontology designed by Barry Smith and his colleagues, which supports information analysis, retrieval and integration across multiple domains. The definition of ‘organization’ I propose is the following: *an object aggregate that bears some normative structure*.

The class object aggregate is defined in BFO as a “material entity that is made up of a collection of objects and whose parts are exactly exhausted by the objects that form this collection” (Arp, Smith & Spear 2015). This class is particularly useful in order to characterize organizations for two reasons. The first is simply that it captures the fact that organizations have members. The members of the organizations are usually *persons*, but they can also be other *organizations*, as in the case of intergovernmental organizations, such as the United Nations or the European Union. The second reason why the class *object aggregate* is useful for our definition is that it allows us to say that an organization preserves its identity through time in spite of the change of its members.

The *normative structure* of the organization defines the roles of its members and the organization’s overall role. The normative structure is a *role aggregate*, as in BFO we can use “aggregate” not only for material entities, but also for entities such as *roles*, which in BFO are *realizable entities*. This role aggregate has as member parts different roles that will be held by the members of the organization. The normative structure is the output of a *speech act* and it can normally be revised through other speech acts over time. A particularly important subtype of speech act is *document act*, since many modern organizations are created, preserved and changed through the use of documents.

Towards a Social Ontology of Work

RACHEL FRASER

University of Cambridge

ABSTRACT

Lots of work is unpleasant, dull, or demeaning. Responses to this fact have come in two flavours: 'better-work politics' and 'anti-work politics'. The former, broadly speaking, affirms the value of work, seeking to organise labour differently and improve the conditions of workers. This tradition, as Weeks (2011) puts it, 'acclaims' work and aims at the restoration of its dignity in an unalienated form'. The anti-work tradition is more radical: it argues not for the reform of work but for its abolition. Anti-work theorists are often written off as hopelessly utopian and impractical; my project in this paper is to defend them.

The argument comes in two parts. First, I argue that we can interpret anti-work Marxists - for example, Silvia Federici - as making a series of tightly related ontological and political claims, which, in the lexicon of analytic metaphysics, make the anti-work theorist an aspirational eliminativist about work. Second, I argue that we can make sense of, and have reasons to endorse, aspirational eliminativism about work.

To be an eliminativist about Fs, roughly speaking, is to say that our talk of Fs fails to pick out any genuine or interesting kind. So, to be an eliminativist about race is to think, like Appiah (1993), that 'there is nothing in the world that can do all we ask race to do for us'. We can contrast Appiah's position on race with that of Haslanger. Haslanger (2012) maintains that our race-talk really does track politically important features of and kinds in the world. Thus, she is not a straightforward eliminativist. However, she maintains that much of our race-talk succeeds in picking out politically important kinds because of the oppressive structures in which we are embedded: to be white, for Haslanger, is (roughly) to occupy a particular social position in a social structure which marks people with certain bodily features for preferential treatment. Such social structures are bad, and we should aim to dismantle them; but if we dismantle them, our race-talk will fail to track politically important features of the world. Thus, insofar as we aspire to a world without white supremacy and racial injustice, we aspire to be eliminativists with respect to our race-talk. More generally, then, to be an aspirational eliminativist about Fs, is to think that our talk of Fs really does pick out a genuine or interesting kind, but to think too that there being any such kind to be picked out constitutively depends upon social practices which are oppressive, dominating, or unjust.

To be an aspirational eliminativist about work, then, is to think that our work-talk in fact picks out some real and important phenomenon in the world, but that the phenomenon it picks out constitutively depends on oppressive, unjust, or dominating structures. I use the resources of neo-Republicanism to suggest - as a first approximation - that we should understand work as constitutively tied to domination. That is I propose a domination condition on work. My domination condition says that work is socially constructed in the following way: roughly, by performing some activity *a*, an agent performs work only if they perform activity *a* in a context in which they are structurally dominated, where an agent is structurally dominated iff their position in a social network makes them vulnerable to the arbitrary exercise of power. On

my account of power, individual and group agents can exercise arbitrary power, as can social structures and natural phenomena.

It turns out that the domination condition will need to be weakened - it is vulnerable to various important objections. I consider two arguments against the domination condition in particular, the Boyle objection and the Shipwreck objection. The Boyle objection says that people who are not structurally dominated are nevertheless capable of working: think of the gentleman scientist, like Robert Boyle, who worked hard to understand the relation between volume and pressure. The Shipwreck objection draws our attention to contexts in which an agent is socially isolated, and so, plausibly, has no social position, and yet must work or starve. In light of these cases, I suggest an alternative, more 'affect' or 'norm-centred' conception of work, on which an activity which is not embedded in patterns of domination can count as work, so long as the activity is governed by sufficiently many norms which are characteristic of domination-embedded activities.

Even this weakened domination condition imposes only a necessary condition: it falls short of giving anything like an analysis of work. Nevertheless, necessary conditions are a good place to start when, and this one comes with a startling result: anti-work theorists succeed in capturing something significant and under-appreciated about the make-up of social reality.

REFERENCES

- Appiah, K. A. 1993. *In My Father's House*. Oxford: Oxford University Press.
Haslanger, S. 2012. *Resisting Reality*. Oxford: Oxford University Press.
Weeks, K. 2011. *The Problem With Work*. London: Duke University Press.

A New Method for Social Preference Aggregation

KALLE GRILL

Umeå University

ABSTRACT

One of the problems with traditional social choice preference aggregation is that it is not sensitive to what is at stake for different people. In their 2010 “Democracy and Proportionality”, Brighouse and Fleurbaey propose as a remedy that power should be distributed in proportion to people’s stakes in the decision under consideration. They call this the proportionality principle. They say it will address the traditional conflict between liberalism and democracy because liberal rights can be viewed as inspired by principle: If I am the only one affected by some decisions, then only I have a stake and so only I should have influence over the decision. However, there is a problem with this solution to the conflict between democracy and liberalism. The problem appears when a decision affects more than one person. In such cases, the proportionality principle may assign great influence to some person because of her high stakes, yet this person may use her power to affect the interests of others, perhaps with no concern for herself. She may either be altruistic and use her power to benefit others, perhaps against their preference, or sadistic and use her power to harm others.

Consider this example of altruism: A group consists of two types of people – As and Bs. There are slightly more As than Bs. There is a policy that can be applied to the whole group or not at all. The policy restricts an unhealthy but pleasant option that both the As and the Bs frequently utilize. The Bs prefer to keep the option available, because they think this is in their best interest. The As prefer that the option be restricted, because they think this is in the best interest of the Bs. In this example, the proportionality principle allows that the policy be enacted, because it is preferred by sufficiently many with sufficiently much at stake, even though no one prefers herself to be subject to the restriction. From a liberal perspective, it might seem that the As join rank with the paternalistic policy-maker in forcing the restriction on the Bs, against their will, gladly paying the price of having the restriction apply also to themselves. Though this consequence can perhaps be accepted, it is in strong tension with the normative basis for the proportionality principle and its allocation of power in accordance with stakes.

The problem of other-regarding preferences is of course well known, going back to Sen’s (1970) “The Impossibility of a Paretian Liberal” and further clarified by Dworkin’s observation (in *Taking Rights Seriously*) that so called “external preferences” make preference utilitarianism unreasonably inequalitarian. It seems clear that the proportionality principle cannot escape it. Note that it is not sufficient to define the stakes themselves as self-regarding: the problem is that the power awarded by having stakes can be used other-regardingly. And if we should decide that it cannot be so used, it would hardly be power. Power that can only be used to promote one’s own interests seems more like forced promotion of one’s own interests.

Against this background, I propose an alternative way of aggregating preferences in a way sensitive to what is at stake. The idea is essentially to use a method of preference purification

proposed for individuals, by Grill 2015, "Respect for What?", and apply it, possibly with modifications, on a collective level. The starting point for this method is that a preference set should be purified to become informed and coherent, without introducing any new preferences. For Grill, the process defines a person's "true" preferences, which should not include any preferences that the person does not actually have. For us, the process defines the aggregate preferences of a group of individuals (the individual preference sets that make up the input to the aggregate may be "true", actual, idealized or whatever).

The preference purification process is aimed at joint satisfiability of preferences, in the sense that things could be as preferred. In order to achieve this, conflicting preferences must be dropped, and what amounts to a conflict is determined by how the world is, not by the person's beliefs. There are many ways to achieve joint satisfiability of this sort, and so the formula for attaining it is minimal loss, where loss is measured in terms of the number and intensity of preferences that are given up (not counting preferences that are undermined because they are conditional on the satisfaction of a preference that is given up). The intensity of a preference is measured in terms of what sacrifices one is prepared to make in order to have it satisfied. To use this process on a collective level requires that the intensity of preferences is interpersonally comparable.

The process can be straight-forwardly used on a collective by simply treating all preferences as the preferences of one individual and then purify to joint satisfiability with minimal loss. Whatever preferences that remain after the process can then be considered to be the preferences of the collective. This process does not exclude other-regarding preferences. Applied to the group of As and Bs in the example, the process will sanction restriction if the preferences for the restriction are more intense and numerous than the preferences against. This may be unlikely, however, given that the As prefer non-restriction for themselves. It is also possible to exclude other-regarding preferences before the process is run, without conflict with any underlying normative starting points. Another optional modification, egalitarian in spirit, is to adjust the intensity of preferences so that each individual either has the same total amount of intensity to distribute over all her preferences, or has the same level of maximal intensity.

This method of preference aggregation produces a social preference set that is sensitive to stakes via sensitivity to the intensity of preferences.

Explaining Moral Responsibility for Structurally Produced Harms

MATTIAS GUNNEMYR

Lund University

ABSTRACT

It is notoriously difficult to decide who is responsible for structurally produced harms such as global warming, traffic congestions or harmful working conditions for labourers within the global garment industry. Typically, the cause of this kind of harm is both mediated through several different levels of agency and overdetermined at several of these levels. Such harms also have in common that they are the result of a confluence of many actions performed by many agents, individual agents as well as collective ones. For these reasons, many influential accounts of joint, shared or collective responsibility – such as those presented by French (1984); Gilbert (2014); Isaacs (2011); Kutz (2000); Pettit (2007) – cannot take us all the way in assigning moral responsibility for structurally produced harms.

By using and developing a Strawsonian account of moral responsibility, I will argue that there is a close connection between who an agent holds morally responsible for an event and what an agent considers to be the explanation for this event and apply this idea to structurally produced harms (for a similar account, see Björnsson, 2014, forthcoming; Björnsson & Persson, 2012). Using Young (2011)'s example of the global garment industry, I will argue that we might hold for instance the global garment companies morally responsible if we conceive their substandard quality of will to be a part of a significant explanation for harm. Similarly, we might hold the consumers morally responsible if we conceive their substandard quality of will to be a part of a significant explanation for harm.

Most defenders of a Strawsonian model of responsibility proceeds by considering cases where there is an obvious cause of harm that we direct our reactive attitudes towards. For instance *A* failed to pick up *B* at the airport as promised or *A* stepped on *B*'s hand (cf. Franklin, 2013; Strawson, 2008 [1962]; Wallace, 1994). However, as these examples usually goes, it turns out that *A* can make some kind of plea that tend to either mitigate *B*'s reactive attitudes towards *A*, or completely remove *B*'s reactive attitude towards *A*. For instance it turns out that *A* failed to pick up *B* at the airport since *A* was tied to a chair or got stuck in unexpectedly bad traffic. What I find unsatisfying is that there seems to be a story to be told about why *A* appears to be blameworthy for the harm in question to begin with. What about cases where there is no single salient explanation for harm? As Young (2011)'s example of the global garment industry illustrates, there are cases in which the explanation of harm is not obvious. For instance, when someone blames the consumers for the harmful working conditions at the manufactories, someone else might blame the global garment companies for this harm. For this reason Strawsonian theories must be complemented with a story about why *A* appears to be blameworthy for harm to begin with. I suggest we do it along these lines:

An agent *A* hold another agent *P* morally responsible for event *E* to the extent that *A* conceive *P*'s quality of will with respect to *E* to be part of a significant explanation of *E*, but only if *A* considers *P* to be normatively competent with respect to *E*.

This account of moral responsibility clarifies the relation between who we hold morally responsible for certain harm and what we conceive of as the explanation for this harm. It also incorporates the characteristic Strawsonian idea that we must conceive the agents we hold morally responsible as being normatively competent and as manifesting a substandard quality of will in their for the harm in question relevant action (for a discussion on these requirements, cf. Franklin, 2013; McKenna, 1998; Stern, 1974; Strawson, 2008 [1962]; Wallace, 1994; Watson, 1987).

On Young's own model of responsibility for structural injustice, agents who are participating in unjust socio-structural processes are never blameworthy, liable nor guilty for these harms. Given this, the conclusion that we might hold agents morally responsible for structurally produced harms might seem surprising. Even though Young's account has many merits, I will argue that one of its assumptions is mistaken. Her major argument for thinking that agents participating in unjust social processes cannot be held blameworthy, liable or guilty is that it is impossible to trace – to *distribute* – the correct amount of blame, liability or guilt to each responsible agent. However, on a Strawsonian account of moral responsibility such as the one I have suggested, there is not a definite amount of blame to distribute. Rather, we hold a particular agent morally responsible if we conceive this agent's substandard quality of will as manifested in their actions to be a part of a significant explanation of harm.

As a final point, the account of moral responsibility I have suggested can also explain our negative reactive attitudes towards agents who could have done something about for instance the harmful working conditions in the garment manufactories (i.e. towards agents who has failed to shoulder their forward-looking responsibility). In the discourse, such a forward-looking responsibility is usually thought of as a responsibility to form a collective together with other agents; a collective that has the power to ameliorate the harmful situation through collective action (cf. Held, 1970; Isaacs, 2011; May, 1990; Petersson, 2008; Tannsjo, 2007). On the account I offer, if we consider such a failure be a significant explanation for why harm continues to be reproduced, and if a particular agent who could have joined or formed such a collective lacks any excusing condition for not doing this, we blame him for this failure.

REFERENCES

- Byörnsson, G. (2014) "Essentially Shared Obligations" in *Midwest Studies In Philosophy*, 38(1), pp. 103-120.
- Byörnsson, G. (forthcoming) "Explaining (away) the epistemic condition on moral responsibility" in P. Robichaud & J. W. Wieland (Eds.) *Responsibility - The Epistemic Condition*: Oxford University Press.
- Byörnsson, G., & Persson, K. (2012) "The Explanatory Component of Moral Responsibility" in *Noûs*, 46(2), pp. 326-354.

- Franklin, C. E. (2013) "A theory of the normative force of pleas" in *Philosophical Studies*, 163(2), pp. 479-502.
- French, P. (1984) *Collective and Corporate Responsibility*: Columbia University Press.
- Gilbert, M. (2014) *Joint Commitment: How We Make the Social World*. New York, NY: Oxford University Press.
- Held, V. (1970) "Can a random collection of individuals be morally responsible?" in *Journal of Philosophy*, 67(14), pp. 471-481.
- Isaacs, T. (2011) *Moral responsibility in collective contexts*: New York: Oxford University Press.
- Kutz, C. (2000) *Complicity: ethics and law for a collective age*. Cambridge: Cambridge University Press.
- May, L. (1990) "Collective inaction and shared responsibility" in *Noûs*, 24(2), pp. 269-277.
- McKenna, M. (1998) "The Limits of Evil and the Role of Moral Address: A Defense of Strawsonian Compatibilism" in *The Journal of Ethics*, 2(2), pp. 123-142.
- Petersson, B. (2008) "Collective Omissions and Responsibility" in *Philosophical Papers*, 37(2), pp. 243-261.
- Pettit, P. (2007) "Responsibility Incorporated" in *Ethics*, 117(2), pp. 171-201.
- Stern, L. (1974) "Freedom, blame, and moral community" in *The Journal of Philosophy*, 71(3), pp. 72-84.
- Strawson, P. F. (2008 [1962]) "Freedom and Resentment " in *Freedom and Resentment and Other Essays*: Taylor & Francis.
- Tannsjö, T. (2007) "The Myth of Innocence: On Collective Responsibility and Collective Punishment" in *Philosophical Papers*, 36(2), pp. 295-314.
- Wallace, R. J. (1994) *Responsibility and the Moral Sentiments*: Harvard University Press.
- Watson, G. (1987) "Responsibility and the limits of evil: Variations on a Strawsonian theme" in F. Schoeman (Ed.), *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*. Cambridge: Cambridge University Press.
- Young, I. M. (2011) *Responsibility for Justice*. New York: Oxford University Press.

Collective Moral Agency and Identity Over Time

NIELS DE HAAN
University of Vienna

ABSTRACT

Elsewhere I have argued that organizations, such as states and NGO's, may hold duties to collectivize if this enables them to solve a particular moral problem (more effectively). This collectivization duty would require the organizations to take responsive steps towards the formation of what I call a *multi-organizational collective* (MOC). Assuming that only moral agents can hold duties, it is imperative that the organizations qualify as moral agents, otherwise they cannot hold the collectivization duty to begin with. Furthermore, it is imperative that the MOC qualifies as a moral agent, otherwise it cannot incur the collective duty to solve the particular moral problem. According to all accounts of collective moral agency, organizations qualify as collective moral agents. However, this does not seem to be the case for MOC's.

Toni Erskine (Erskine, 2003), building on Peter French (French, 1984), claims that a collective is a candidate for collective moral agency if it has the following: an identity that is more than the sum of the identities of its constitutive parts and, therefore, does not rely on determinate membership; a decision-making structure; an identity over time; and a conception of itself as a unit. In this article the locus is on one condition in particular, namely identity over time. This criterion is invoked in order to exclude groups that are transient or spurious from collective moral agency. Collective moral group agents must have continuity and according to Erskine this means that they must have a past accessible to experience-memory and a future accessible to intention. In this article I argue that having an identity over time ought not to be a necessary condition for collective moral agency.

While relating retrospective collective moral responsibility to (the failure of) the formation of collectives is nothing new (Held, 1970), it has only been more recent that attention is being paid to prospective collective moral responsibility and the formation of collectives (Collins, 2013). In the latter case the idea is that agents ought to form collective moral agents who incur a collective duty to solve a collective action problem that is morally pressing. I call these resulting collective moral agents from collectivization duties, on both the individual and collective level, *duty-bound collectives*. This might explain why according to Erskine's account newly formed collectives do not seem to qualify as collective moral agents. There is a significant difference between becoming an (full-fledged) individual moral agent and a collective moral agent. Once an individual becomes a moral person (in the fullest sense), she already has a considerable past available to her. This is not necessarily the case for collective agents. Newly formed collectives, including duty-bound collectives, seem to have a blank slate, a *tabulae rasae* if you will. Therefore, they do not have an identity over time, thus do not qualify as collective moral agents. One cannot solve this by making reference to the pasts of the collective's members, since the identity of the collective ought to be more than the sum of the identities of its constitutive parts.

However, I argue that the criterion of having an identity over time is too stringent for collec-

tive moral agency. I present a hypothetical yet not impossible case of retrograde amnesia where an agent her episodic memory is severely affected while her semantic memory remains intact in the relevant way. This means that the agent may continue to remember general knowledge, including a moral compass, but fail to remember any specific events in her life. The agent is still in possession of the relevant degree of rationality and autonomy. The agent commits both a praiseworthy and blameworthy act. In both cases we would continue to hold the person responsible for what she did and consider her a moral agent. This case is meant to show that *actually* having an identity over time is not a necessary condition for individual moral agency.

Since we ordinarily do not impose more stringent conditions on collective moral agency than we do on individual moral agency, I argue that it would be inconsistent to invoke having an identity over time as a necessary condition for collective moral agency. I further show that duty-bound collectives do not have a troublesome relation with the kind of continuity that the criterion of having an identity over time should foster.

Next, I consider whether the *capacity* to have an identity over time should be a necessary condition for collective moral agency. The problem is that by itself the condition does not perform its intended function, namely excluding spurious or transient collectivities from moral agency. However, I argue that advocates of a stringent account of collective moral agency are not at a loss, as the range of groups considered to be collective moral agents can equally well be (justifiably) determined by other criteria.

Finally, I reflect on a more general worry concerning accounts of collective moral agency that is highlighted by this discussion. Generally, most accounts present ideal cases of group agents to show that it would be justifiable to consider them moral agents. But we do not require individual moral agents to be rational all the time, they simply need the capacity to do so and exercise that capacity when appropriate. Therefore, it seems incoherent and too stringent to require collective moral agents to be rational all the time. Similarly, they only need to have the capacity to be rational. This should be reflected in the conditions we invoke for collective moral agency. The question, then, becomes how we define this capacity without expanding the range of moral agents too broadly.

Concluding, since *actually* having an identity over time is not a necessary condition for collective moral agency, newly formed collectives, including duty-bound collectives, may qualify as collective moral agents. This shows that when one is devising an account of collective moral agency, it is important that one does not just consider existing collectives, but the to-be-formed-collectives as well, as this could have important ramifications in the field of (global) ethics.

Causal Powers and Social Ontology

TOBIAS HANSSON WAHLBERG

Lund University

ABSTRACT

Since the early 1970s, non-Humean causal powers – often referred to simply as *powers* – have made a prominent return to philosophy. Initially invoked primarily in the philosophy of physics and chemistry (e.g. Harré and Madden 1975; Bhaskar 1975/2008; Cartwright 1983), these powers were subsequently put to use in metaphysics (e.g. Shoemaker 1980; Ellis 2001), the philosophy of biology (Dupré 1996), and, in recent years, the philosophy of social science and social ontology (e.g. Archer 1995; Cartwright 1999; Elder-Vass 2010; Groff 2011; Kaidesoja 2013; Lawson 2013; but see already Bhaskar 1979/1998). Indeed, there is now an interdisciplinary movement, called critical realism, which among other things applies powers to the social realm. In this paper, I question the deployment of powers within the social domain. I do so by arguing that alleged social powers are either contradictory or causally redundant entities. I also offer a diagnosis of how and where defenders of that deployment go astray. I do so by distinguishing between true ability predications and powers, between constitutive (or “performative”) abilities and powers, and between deontic and causal powers. These important distinctions tend to be jumbled up in the relevant literature.

REFERENCES

- Archer M., 1995, *Realist Social Theory: The Morphogenetic Approach*, Cambridge: Cambridge University Press.
- Bhaskar R., 1975/2008, *A Realist Theory of Science*, London: Verso.
- Bhaskar R., 1979/1998, *The Possibility of Naturalism: A Philosophical Critique of the Contemporary Human Sciences*, London: Routledge.
- Cartwright N., 1983, *How the Laws of Physics Lie*, Oxford: Oxford University Press.
- Cartwright N., 1999, *The Dappled World*, Cambridge: Cambridge University Press.
- Dupré J., 1993, *The Disorder of Things*, Cambridge, MA: Harvard University Press.
- Elder-Vass D., 2010, *The Causal Power of Social Structures*, Cambridge: Cambridge University Press.
- Ellis B., 2001, *Scientific Essentialism*, Cambridge: Cambridge University Press.
- Groff R., 2011, “Getting Past Hume in the Philosophy of Social Science”, in McKay P., Russo F. And Williamson J. (eds.), *Causality in the Sciences*, Oxford Scholarship Online.
- Harré R. and Madden E. H., 1975, *Causal Powers: A Theory of Natural Necessity*, Oxford: Basil Blackwell.
- Kaidesoja T., 2013, *Naturalizing Critical Realist Social Ontology*, London: Routledge.
- Lawson T., 2013, “Emergence and Social Causation”, in Groff R. and Greco G. (eds.), *Powers and Capacities in Philosophy: The New Aristotelianism*, 2013, New York: Routledge.
- Shoemaker S., 1980, “Causality and Properties”, in P. van Inwagen (ed.), *Time and Cause: Essays Presented to Richard Taylor*, pp. 109-35, Dordrecht: Reidel

What's the Difference Between Money and Gender? Social Construction, Critique, and Change

FRANK HINDRIKS

University of Groningen

ABSTRACT

When someone claims that gender is socially constructed, she will in all likelihood be critical of gender in a way that someone who claims that money is socially constructed will typically not be. The point of a critical social construction claim is that, in spite of appearances, change is possible as well as desirable (Hacking 1999). The desirability of change could be part of the very meaning of social construction claims. I argue instead that their meaning is not evaluative in this way.

Instead, it concerns the fact that social constructs depend on human attitudes. Such attitude-dependence captures the core of the ontology of social constructs, which gender and money have in common. It entails that social constructs are to some extent subject to change. In contrast to money, people are typically ignorant about the fact that gender is a social construct. Hence, they rarely realize that it can be changed. I go on to argue that the desirability of change is a conversational implicature that social construction claims have in particular contexts (Grice 1975).

The striking thesis that surfaces from this analysis is that the ontology of money and gender is basically identical. The differences between them reside in other aspects of these social constructs – conceptual, pragmatic, epistemic, and normative – and can be understood from within a unified framework of social construction.

How Do Non-Joint Commitments Come Into Being? : An Attempt at Cultural Naturalism

INGVAR JOHANSSON

Umeå University

ABSTRACT

I will question the view that social ontology *always* can take an autonomous I-mode of individuals (a self-consciousness) for granted, and that, therefore, only the creation of we-modes are in need of analysis. Empirical evidence clearly points towards both the developmental-psychological fact that children are conscious before they become self-conscious, and the historical-anthropological fact that individualism arises late in the history of mankind. A comprehensive social ontology must be able to make sense of the structure beneath both the emergence of individual self-consciousness and social individualism.

My proposal is that the problems hinted at can be dealt with if a distinction between *reflective* and *unreflective* (or *pre-reflective*) conscious events/acts/states with intentionality is introduced. This is so since even unreflective intentional states allow for a distinction between the I-mode and we-modes. For example, an individual can have a joint or shared commitment, a commitment in some we-mode or nested intentionality, without being reflectively aware of it. (Some of the views to be put forward can be found in my paper "How Do Non-Joint Commitments Come Into Being? An Attempt at Cultural Naturalism", in Psarros and Scholte-Ostermann, eds., *Facets of Sociality*, 2006, pp. 135–49. See <http://www.ingvarjohansson.se> Ontology/On Social Ontology.)

The Explanatory Power of the Axiological Objects of Collective Intention

TAINA KALLIOKOSKI

University of Helsinki

ABSTRACT

Many scholars in the field of social ontology discuss and inquire the gap between or transition from individual intentionality to collective intentionality. Not so many are interested in the question, what the reasons and conditions are under which the collective intentionality in voluntary cooperating situations maintains or dissolves.

In this paper, I argue that we need to take the axiological objects of intentions into account if we want to understand why social groups continue to cooperate or cease to exist.

The philosopher G. H. von Wright (1968) pointed out that an agent rarely intends to pursue bare factual objects such as the changed state of affairs but axiological objects as well: the agent intends that the results and consequences of his actions are valued at least by herself and by those to whom the action affects. The intention of the action may succeed or fail in regard to its factual or axiological object.

The distinction between the factual and the axiological object of intention is crucial also when we explain collective intentions in a voluntarily organized group. The distinction may help to understand the reasons why every social good pursued by cooperation is not achieved or why the social associations dissolve. I will enlighten the explanatory power of the axiological object of collective intentions with analyzing a cooperation example, Hoffice co-working. In explaining collective action, I use R. Tuomela's account of the weak I-mode collective intentionality. (Tuomela 2007 and 2013.)

Hoffice is an example of the modern voluntary cooperation of people who work with independent projects, usually without full-time employments, e.g. students, freelancers, and researchers. Hoffice combines the words, places, and atmospheres of 'home' and 'office'. A person who participates in Hoffice usually wants to strengthen her self-control and create a sense of belonging. The idea of Hoffice is to work with others in structured sessions and in an informal manner by assembling in the same location at the same time. Participants announce their private goals for the working session and promote their individual projects through the sessions of the day. Afterward, they reflect the quality and accomplishments of the session by stating aloud how their projects have progressed and intentions materialized. (Toikkanen 2016.)

In Tuomela's terms, Hoffice is knit together with the weak collective intention of the participants, thus it is an example of the pro-group I-mode action. Every participant intends to cooperate, but they do so for private reasons. They also intend to pursue their private goals such as enhancing their projects and gaining self-control through the social pressure and the support of others. The goals are private but the means to pursue the private goal is the same for everyone: co-working. By applying Tuomela's idea that the I-mode cooperation has both teleological and social content (2007, 151–152), we can say that the Hoffice participants' teleological content is to get their projects advanced and the social content covers all the social interaction

that happens during the Hoffice meeting and enables the co-working. Participants may have factual and axiological intentions in relation to the both teleological and social content.

Hoffice participants share the factual object of teleological intention: The factual object of the shared intention is to co-work in the same place at the same time in an agreed manner. After the Hoffice day, the telos of cooperation has been accomplished if the participants have worked through the day. In the pro group I-mode, the axiological objects of the teleological intention are both shared by the participants and private. Each participant intends to gain the beneficial or good consequences of the collective action in the private I-mode. One shared value in every Hoffice co-working group is the cooperation itself: Hoffice group intends the co-operation style to be valued as beneficial by and for cooperating agents, otherwise they would not have had reasons to participate. If a participant makes a judgment that the co-working has failed in regard to its axiological object, the motivating reasons to participate and continue cooperation may cease to exist.

In voluntary I-mode cooperation, the participants are both evaluating subjects and experiencers of the consequences of cooperation: they assess the beneficial and harmful effects comparing them with the axiological objects of their intention. The group falters, if the cooperation is criticized or not valued afterward because both the motivating reason to participate and the realm of concern of Hoffice groups are based on the appreciation of the specific style of cooperation.

The continuity of voluntary I-mode groups requires a constant process of the positive axiological judgments of cooperation. The judgments can be based on the evaluation and interpretation of experiences attained during the cooperation, ethical values, aesthetic values or hedonistic preferences. If they are not positive, the motivating reasons for the agent in the I-mode to cooperate diminish.

REFERENCES

- Toikkanen, Jarkko. 2016. "Home Office (Hoffice) – En etnografisk studie om ett coworking-fenomen." Avhandling. Institutionen för företagsledning och organisation. Hanken Svenska handelshögskolan. Helsingfors. <https://helda.helsinki.fi/dhanken/bitstream/handle/10138/163756/toikkanen.pdf>
- Tuomela, Raimo. 2013. *Social Ontology: Collective Intentionality and Group Agents*. New York: Oxford University Press.
- Tuomela, Raimo. 2007. *The Philosophy of Sociality: A Shared Point of View*. New York: Oxford University Press.
- von Wright, Georg Henrik. 1968. *Varieties of Goodness*. London: Routledge & Kegan Paul.

Declarations and institutional reality

ARTO LAITINEN

University of Tampere

ABSTRACT

This paper examines the nature of declarations, in order to assess Searle's claim:

With the important exception of language itself, all of institutional reality ... is created by speech acts that have the same logical form as Declarations. Not all of them are, strictly speaking, Declarations, because sometimes we just linguistically treat or describe, or refer to, or talk about, or even think about an object in a way that creates a reality by representing that reality as created."(Searle 2010, 12-13).

The paper distinguishes "ritual" and "non-ritual" contexts and argues against the centrality of the "logical form of declarations" in both. In ritual contexts, such as weddings, or adjourning a meeting, uttering a codeword, or performing an act of suitable kind, is needed in order to make a change in institutional reality. Whether or not the codeword "represents the reality as existing" is beside the point: whether the practice is that the chairperson utters the words "The meeting is adjourned" or, say, "abracadabra", does not change the logic of how institutional acts are made. Even when the utterances "represent the reality as existing" they do not create any reality because they do so; but because they count as codewords of the required kind.

Outside such institutionalized practices or ritual contexts, it could be that declarations are a subclass of speech acts (alongside assertives, commissives, directives, expressives), needed in the genesis of institutions. The paper presents an argument that Searle lacks the resources to distinguish declarations from the other types of speech acts (and indeed, Searle mentions promises, orders, and apologies as examples of both declarations and other speech act types). All speech acts may include the relevant performative aspect.

The conclusion of the paper is that the version of Searle's account which focuses on collective acceptance and constitutive rules is superior to the version which focuses on declarations. (The problem of how these two versions are meant to co-exist in Searle's 2010 book has been discussed earlier e.g. By Tsohatsidis 2010).

Social Construction and Universality

ESA DIAZ LEON

University of Barcelona

ABSTRACT

What does it mean to say that a category is socially constructed? In my view, it is not very useful to look for the notion of social construction; rather, there are different notions of social construction, for different purposes. Therefore, a more promising approach is to ask the question of which project is at issue in some particular context, and which notion of social construction is more useful for the purposes of that project. In previous work, I have focused on the project of arguing against the inevitability of a trait, and allowing the possibility of social change. In this paper, my aim is to focus on the project of arguing against the universality of a trait. I will examine several notions of social construction in the literature, and I will ask, for each of them, whether it entails that if X is socially constructed in that sense, then X is not universal.

In order to answer this question, I will distinguish three different notions of universality, namely, (i) universal in the sense of being universally instantiated, (ii) universal in the sense of appearing in law-like generalizations (i.e. being an objective or deep property), and (iii) universal in the sense of being independent from some particular cultural or social practices (i.e. being trans-cultural, or culture-independent).

Moreover, I will examine four different notions of social construction (and four corresponding versions of social constructionism), to wit: causal vs. constitutive social constructionism, empty-category constructionism, and superficial constructionism. These different notions of social construction can be understood as follows:

- A property is constitutively socially constructed just in case it is part of the definition (or the nature) of the property that in order to instantiate it, one must instantiate a certain social role, whereas a property is causally socially constructed just in case certain social factors play a causal role in bringing about that feature of individuals.
- Empty-category constructionism: a property is empty when it is not instantiated in the actual world.
- Superficial constructionism: A property is superficial, or not objective, just in case it does not appear in law-like generalizations.

In addition, I draw the following further distinction within the notion of constitutive social construction:

- Culture-dependent constructionism: A subject S has property X only when S instantiates certain social or cultural facts (or enters in certain social or cultural relations).
- Concept-dependent constructionism: A subject S has property X only when someone applies the concept of X to them.

My main aim in the paper then is to assess whether these different versions of social con-

structionism entail or not that a property X is universal in those different senses above. My central conclusions will include the following: On the one hand, I will argue that if a property is constitutively socially constructed, it is clearly culture-dependent, but the claim that a property is causally socially constructed does not always have that consequence. On the other hand, I will argue that if concept-dependent constructionism is true about property X, then no-one can have property X before the concept of X was introduced, but the claim that culture-dependent constructionism is true about X does not always have this consequence. In addition, I will argue that the claim that a property is culture-dependent does not entail that it is superficial or not objective, and vice-versa: the claim that a property is superficial does not entail that it is culture-dependent. That is to say, there can be properties that are culture-dependent but objective (such as social kinds like recession or inflation), and there can be properties that are superficial but culture-independent (such as sleeping on one's back ("backer"), or one's front ("fronter")).

How to Think About Constructions: Towards an Account of the Acquisition of Social Kind Concepts

ROBIN GUID LOEHR

Berlin School of Mind and Brain

ABSTRACT

Abstract I propose an account of concept acquisition of socially constructed kinds like *race*, *racism*, *money* or *president*. This project, for the first time, brings together three different literatures, the literatures of metaphysics of social kinds, reference and concept acquisition (note that the latter usually focus on concrete natural kind concepts). I propose a theory of acquisition for three kinds of social kinds as recently proposed by Khalidi (2015). My main conclusion will that acquiring a social kind concept is not particularly more difficult than acquiring natural kind concepts. This tells us something interesting about the nature of social kinds in general and our relationship to them.

INTRODUCTION

Many of the categories we care most deeply about are social, meaning they do not capture structures in the natural world, but something that exists only in our social world. Uncontroversial examples are *president of the United States of America*, *bachelor*, *money* or *tax payer*. More controversial examples are *race*, *gender* or *truth*. In this paper I am interested in the question of how we are able to think about kinds if they are constructed, i.e., if they are not biologically real.

The most natural way to think about how we can refer to social categories in thought and language is that we represent descriptions that allow us to distinguish the classes we are interested in from those we decide to ignore. For example a plausible theory of how we are able to form a belief about bachelors is that we *know* what distinguishes the category of bachelor from other categories, e.g., by knowing that bachelors are essentially unmarried men. If I decide to think not about bachelors but about unmarried women I can do so by changing the description or definition. I can also give this category a name, say 'bachelorettes'.

Unfortunately, descriptivist views of conceptual and linguistic content face a number of serious problems. First, we often seem to have difficulties defining concepts. This seems to be especially difficult for many social categories like democracy, race or gender. This is a problem not only for philosophers interested in conceptual analysis, but especially for everyone who want to explain the use of concepts and words of lay people. For if we do not seem to know what essentially distinguishes women from men (as we currently do not), how can this knowledge be responsible for our ability to think or talk about gender?

Secondly, the descriptions we do associate with the respective kind are sometimes false, which does not seem to affect our ability to refer to them. This is again especially the case for constructed kinds. One example for such an error in our supposedly reference-fixing descrip-

tions is that people often believe that gender or race are natural categories. The term ‘race’ for example has been defined as “a sort of human group whose members have a distinguishing essence” (Mallon, 2015, 3). Most contemporary biologists however would claim that such an essence does not exist (Sterelny & Griffiths, 1999; Hall, 1996; Taylor, 2013). Does this mean the content of many social terms and concepts do not refer?

A third serious problem is that people often have fundamentally different experiences especially with social categories. Consequently, they will associate or infer different knowledge from the respective category. For example, while republicans and democrats will probably make similar experiences with natural kind objects like gold or dogs, they may associate fundamentally different beliefs with the category *race* despite picking out the same entity. However, we do not usually have the impression that when democrats and republicans have debates on race that they are debating about different things.

These problems are reminiscent of the familiar descriptivist- anti-descriptivist debate associated with Kripke (1972) Putnam (1973) and Burge (1979). Kripke and Putnam argued that proper names and natural kind terms may refer even if we have erroneous or limited knowledge about their referents. Instead of descriptions, proper names are assigned to an entity or category in the world, say a person or a natural kind and gets transmitted from person to person in a causal-historical manner. In this case, even if two people associate vastly different knowledge with a name, they could still be said to refer to the same object if in the same causal-historical relation with that object.

More recently, Boyd (1992), Haslanger (2003, 2005, 2012) and others have applied this theory of reference to social categories. They argue that the referents of our social words are determined not by descriptions, but externally by some causal-historical link. This solution is extremely attractive as it does not require that we have many true beliefs about categories like race and gender and makes due to the intuition that we can make discoveries in the social realm. However, social anti-descriptivist views do not come without problems. The main problem is that it is not clear how we can establish and sustain a non descriptivist connection to something that does not exist in nature, e.g., that we cannot simply ostensively name.

Such problems were recently discussed by Ron Mallon (2015) by proposing a more straight-forward solution to what he calls *the problem of the missing referent and the qua problem*. In this paper I built on this work and propose a theory of concept acquisition in a non-descriptivist framework based on Fodor (1998), Margolis (1998) and Laurence & Margolis (2002, 2011). In particular I argue that socially constructed kinds do not pose any particular difficult problem to anti-descriptivist theories of content because if Mallon (2003; 2015) is right there might be several crucial similarities between natural kind and social kind term that might help to explain how we can think about constructed kinds.

REFERENCES

- Bloom, P. (2000). *How children learn the meanings of words* (Vol. 377). Cambridge, MA: MIT press.
- Burge, T. (1979). Individualism and the Mental. *Midwest studies in philosophy*, 4(1), 73-121.
- Cain, M. J. (2013). Learning, Concept Acquisition and Psychological Essentialism. *Review of Philosophy and Psychology*, 4(4), 577-598.

- Carey, S. (2009). *The origin of concepts*. Oxford University Press.
- Epstein, B. (2008). The realpolitik of reference. *Pacific Philosophical Quarterly*, 89(1), 1-20.
- Evans, G., & McDowell, J. H. (1982). *The varieties of reference* (pp. 100-ff). J. McDowell (Ed.). Oxford: Clarendon Press.
- Evans, G., & Altham, J. E. J. (1973). The causal theory of names. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 47, 187-225.
- Gelman, S. A. (2003). *The Essential Child: Origins of Essentialism in Everyday Thought*. New York, Oxford University Press.
- Griffiths, P. E. (2004). Emotions as natural and normative kinds. *Philosophy of science*, 71(5), 901-911.
- Hacking, I. (1995). The looping effects of human kinds. *Causal cognition: A multidisciplinary debate*, 351-383.
- Hacking, I. (1999). *The social construction of what?*. Harvard university press.
- Hall, S. (1996). New ethnicities. *Stuart Hall: Critical dialogues in cultural studies*, 441-449.
- Haslanger, S. (2005). What are we talking about? The semantics and politics of social kinds. *Hypatia*, 20(4), 10-26.
- Khalidi, M. A. (2015). Three kinds of social kinds. *Philosophy and Phenomenological Research*, 90(1), 96-112.
- Keil, F. C. (1989). *Concepts, Kinds, and Cognitive Development*. Cambridge, Mass., MIT Press.
- Kripke, S. A. (1972). *Naming and necessity*. Springer.
- Laurence, S., & Margolis, E. (2002). Radical concept nativism. *Cognition*, 86(1), 25-55.
- Putnam, H. (1975). The meaning of 'meaning'. In Pessing, A, G. S., editor, *The Twin Earth Chronicles: Twenty Years of Reflection on Hilary Putnam's The meaning of meaning*, pages 3– 52. London: Routledge.
- Machery, E. (2014). "Social Ontology and the Objection from Reification." *Perspectives on Social Ontology and Social Cognition*. M. Gallotti and J. Michael. Dordrecht, Springer: 87–102.
- Mallon, R. (2015). Social Construction and Achieving Reference. *Noûs*.
- Mallon, R. (2003). Social construction, social roles, and stability. In F. Schmitt (ed.), *Socializing Metaphysics : The Nature of Social Reality*. Rowman & Littlefield, pp. 327—54.
- Margolis, E. (1998). How to acquire a concept. *Mind & Language*, 13(3), 347-369.
- Margolis, E., & Laurence, S. (2011). Learning matters: The role of learning in concept acquisition. *Mind & Language*, 26(5), 507-539.
- Medin, D. and Ortony, A. 1989: Psychological Essentialism. In S. Vosniadou and A. Ortony (eds), *Similarity and Analogical Reasoning*. Cambridge University Press, 179–95.
- Mayr, E. (1984). "Species Concepts and Their Application." *Conceptual Issues in Evolutionary Biology: An Anthology*. E. Sober. Cambridge, MA, MIT Press: 531–540.
- Searle, J. R. (1995). *The construction of social reality*. Simon and Schuster.
- Soja, N., Carey, S., & Spelke, E. (1991). Ontological categories guide young children's inductions on word meaning: object terms and substance terms. *Cognition*, 38, 179–211.
- Sterelny, K. and P. E. Griffiths (1999). *Sex and Death: An Introduction to the Philosophy of Biology*. Chicago, Ill., University of Chicago Press.

Heuristics, bounded rationality, and joint action

JUDITH MARTENS

Ruhr Universität Bochum

ABSTRACT

Two dichotomies and their relation seem to lie at the heart of the debate on action and joint action: one between mere bodily movement and (intentional) action, the other between emergent coordination and planned coordination. I argue that a focus on these dichotomies leaves us wanting for “something in between”. I investigate this ‘in between’ based on literature on heuristics (ecological or social rationality). Heuristics trigger purposeful behavior, which seems to involve decisions (they are not mere bodily movements) that are however not structured by explicitly available attitudes.

I want to follow Bratman’s conceptualization of human agents as planning agents. Such agents can share pro-attitudes, which is a possible way to act together. I say possible because Bratman does not argue that this is the only way that agents can share intentions. In several books and articles Bratman emphasizes that his account is meant to demystify complex forms of agency. To do so, he uses creature constructions, showing how we can understand such complex structures through the use of simpler building blocks. This approach – where human agency is modeled according to core features of our acting and our cognitive capacities and limitations – balances between “showing complex agency to be possible” and “describing actual functioning”. The first part of my talk will focus on Bratman’s creature construction in collective intentionality. I will focus on two aspects specifically, (1) on the way the theory deals with limitations to our agency and rationality, and (2) on the specific way Bratman characterizes his way to shared intentionality.

Bratman argues that, due to time constraints and limited cognitive and conative capacities, planning agents often act based on policies and dispositions. This is, however, not translated to cases of shared intentionality. Assuming that we – together – act on such policies and dispositions opens up new ways to understand joint action and collective intentionality. Yet, it also poses some problems, because certain constraints that are built into Bratman’s conception of shared intentions seem non-translatable to joint action based on policies and dispositions.

I will explore an understanding of actions based on policies and dispositions by looking at ecological rationality. First I will introduce the idea of ecological rationality and how it can structure and shape interactions, exploring whether we can understand interaction and joint action based on this form of rationality. How could heuristics structure joint actions? I will take the tit-for-tat heuristic, imitate the successful heuristics, and imitate the majority heuristic as case studies. Secondly, I will compare such an approach based on heuristic and ecological rationality with Bratman’s theory on policies and dispositions. One major difference between heuristics and Bratman’s theory is the boundedness of the rationality.

This paper is an attempt to see whether Bratman’s theory on planning agency and shared (planning) agency is combinable with ideas on bounded rationality. I argue that the com-

bination forces us to be clearer about the status of Bratman's theory: is it purely meant as demystification, or is it an attempt at understanding human agency in everyday life? Heuristics are cognitively less demanding and in that sense in line with Bratman's proposal for policies and dispositions. Yet, heuristics are characterized by rationality that is only consistent within a certain situation, where Bratman's proposal on policies and dispositions are placed within his broader hierarchical model of agency with overall consistency amongst beliefs, desires, intentions and plans. I will argue that the two are less closely connected than suggested in Bratman's theory, but that we should still take the impact of plans on bounded rationality seriously. At the same time, bounded rationality heavily depends on the situation. Multiple agents can share a specific situation, which might serve similar heuristic rationality within each agent. An understanding of joint action and collective intentionality through both shared intentions and heuristics would provide us with more ways to demystify how humans are so capable when it comes to acting together.

Collective memory: Metaphor or reality?

KOURKEN MICHAELIAN

University of Otago

ABSTRACT

The concept of collective memory has long had currency both in the social sciences, with research focusing on how societies and other large-scale groups remember their pasts (Connerton 1989, Misztal 2003), and in psychology, with research looking at how married couples (Harris et al. 2011), mother-child dyads (Reese et al. 1993), and other small-scale groups remember events together (Barnier et al. 2008, Michaelian and Sutton 201x). With few exceptions (Wilson 2005, Sutton 2008), however, philosophers have so far paid it little attention. Consequently, neither philosophical analyses of remembering nor philosophical theories of collective phenomena have yet been brought to bear on collective memory, and it remains unclear just how seriously we should take the concept: is collective memory a mere metaphor, or are (some) groups literally capable of remembering?

This question has two aspects. First, *collectivity*: is collective memory really collective? Second, *mnemicity*: is collective memory really *memory*? Building on theories of collective intentionality and analyses of the concept of memory, the talk will defend a positive answer to the first aspect of the question but a negative answer to the second.

Collectivity: Beginning with the first aspect, the talk explores the capacity of a range of concepts from the collective intentionality literature to capture a sense in which collective memory might be robustly collective. The focus of this part of the talk is primarily on the sort of small-scale groups studied in psychology. A natural starting point is the idea that the collectivity of collective memory can be understood in terms of *collective belief* (Gilbert 1989) or acceptance (Tuomela 2000). While natural, this idea is difficult to reconcile with the fact that collective remembering may involve conflict among group members, as well as the fact that a group need not settle on a determinate representation of the past before it can be said to remember. Another idea is that collective memory can be understood in terms of *joint attention* to the past (Hoerl and McCormack 2005). This idea is more promising but may ultimately be applicable only to certain specific forms of collective memory. A third idea—that collective remembering can be understood as a form of *joint action*—is therefore worth exploring in detail. Accounts of joint action as resulting from joint intention (Searle 1990, Bratman 2014) arguably fail to capture both the fact that collective memory typically does not result from joint intention and the fact that it involves ongoing interaction among group members. An account of joint action in terms of alignment systems (Tollefsen et al. 2013) better captures these facts; such an account, moreover, fits well with recent work arguing that groups sufficiently close-knit to qualify as *transactive memory systems* in Wegner's sense (Wegner 1987) are capable of remembering more or different features of events than their members, thus manifesting a form of emergent group memory (Theiner 2013, Huebner 2016, Harris et al. 201x). Overall, then, while more work remains to be done to establish the collectivity of memory in large-scale groups, a case can be made that memory is robustly collective at least in small-scale groups.

Mnemicity: Turning to the second aspect of the question, the talk looks at whether we can

identify a sense in which collective memory is robustly mnemonic. The focus in this part of the talk is primarily on the sort of large-scale groups studied in the social sciences. Analyses of remembering adopt either of two basic perspectives (Michaelian 2016, Robins 201x). The *encoding-consolidation-storage-retrieval* (E-C-S-R) perspective views remembering as a matter of preserving information over time. The *mental time travel* (MTT) perspective, in contrast, views remembering as a form of constructive or imaginative mental time travel into the past. There are important differences between these perspectives, but, regardless of which is adopted, the alleged analogy between individual and collective remembering quickly breaks down. If the E-C-S-R perspective is adopted, we should expect to find patterns of remembering and forgetting at the collective level that are at least roughly analogous to those found at the individual level. This is precisely what Anastasio et al. (2012) and Tanesini (201x) claim to find, suggesting, for example, that interruptions to the consolidation process lead to *collective amnesia*, just as they lead to individual amnesia. Upon closer inspection, however, this appears not to be the case, as the patterns of remembering and forgetting at issue in collective “amnesia” differ radically from those at issue in individual amnesia; for example, collective consolidation can be resumed at a later date if interrupted, but individual consolidation cannot. If the MTT perspective is adopted, we should expect to find that groups engage in mental time travel not only into the past but also into the future. This is precisely what Szpunar and Szpunar (2016) claim to find, suggesting that groups engage in future-oriented mental time travel much in the way that individuals engage in future-oriented mental time travel. Upon closer inspection, however, the most plausible examples of collective future-oriented mental time travel turn out to be cases of individual *future-oriented mental time travel* in the context of a group; moreover, individual mental time travel is characterized by phenomenological features which likely have no analogues at the collective level. Overall, then, while more work remains to be done to establish that collective memory in small-scale groups is not, strictly speaking, mnemonic, a case can be made that at least large groups are incapable of remembering in an interestingly robust sense, i.e., that the collective phenomena identified in the first section do are not, strictly speaking, instances of collective *memory*.

The talk ends with a brief discussion of the implications of the positive and negative aspects of its conclusion for the interdisciplinary field of memory studies and a call for increased attention to collective memory on the part of social ontologists.

REFERENCES

- Anastasio, T.J., K.A. Ehrenberger, P. Watson, and W. Zhang. 2012. *Individual and Collective Memory Consolidation: Analogous Processes on Different Levels*. MIT Press.
- Barnier, A., J. Sutton, C. Harris, and R. Wilson. 2008. A conceptual and empirical framework for the social distribution of cognition: The case of memory. *Cognitive Systems Research* 9(1-2):33-51.
- Bratman, M. E. 2014. *Shared Agency: A Planning Theory of Acting Together*. Oxford University Press.
- Connerton, P. 1989. *How Societies Remember*. Cambridge University Press.
- Gilbert, M. 1989. *On Social Facts*. Routledge.
- Harris, C. B., A. J. Barnier, J. Sutton, P. G. Keil, and R. A. Dixon. 201x. “Going episodic”:

Collaborative inhibition and facilitation when long-married couples remember together. *Memory*.

- Harris, C. B., P. G. Keil, J. Sutton, A. J. Barnier, and D. J. F. Mcllwain. 2011. We remember, we forget: Collaborative remembering in older couples. *Discourse Processes* 48(4):267-303.
- Hoerl, C. and T. McCormack. 2005. Joint reminiscing as joint attention to the past. *Joint Attention: Communication and Other Minds*. Eds. N. Eilan, C. Hoerl, T. McCormack, and J. Roessler. Pp. 260–286. Oxford University Press.
- Huebner, B. 2016. Transactive memory reconstructed: Rethinking Wegner's research program. *Southern Journal of Philosophy* 54(1):48-69.
- Michaelian, K. 2016. *Mental Time Travel: Episodic Memory and Our Knowledge of the Personal Past*. MIT Press.
- Michaelian, K., and J. Sutton. 201x. Collective memory. *The Routledge Handbook of Collective Intentionality*. Eds. K. Ludwig and M. Jankovic. Routledge.
- Misztal, B. 2003. *Theories of Social Remembering*. McGraw-Hill.
- Reese, E., C. A. Haden, and R. Fivush. 1993. Mother-child conversations about the past: Relationships of style and memory over time. *Cognitive Development* 8(4):403-430.
- Searle, J. R. 1995. *The Construction of Social Reality*. Simon and Schuster.
- Robins, S. K. 201x. Confabulation and constructive memory. *Synthese*.
- Sutton, J. 2008. Between individual and collective memory: Coordination, interaction, distribution. *Social Research* 75(1):23-48.
- Szpunar, P. M. and K. K. Szpunar. 2016. Collective future thought: Concept, function, and implications for collective memory studies. *Memory Studies* 9(4):376-389.
- Tanesini, A. 201x. Collective amnesia and epistemic injustice. *Socially Extended Epistemology*. Eds. J. A. Carter, A. Clark, J. Kallestrup, S. O. Palermos, and D. Pritchard. Oxford University Press.
- Theiner, G. 2013. Transactive memory systems: A mechanistic analysis of emergent group memory. *Review of Philosophy and Psychology* 4(1):65-89.
- Tollefsen, D. P., R. Dale, and A. Paxton. 2013. Alignment, transactive memory, and collective cognitive systems. *Review of Philosophy and Psychology* 4(1):49–64.
- Tuomela, R. 2000. *Cooperation: A Philosophical Study*. Springer.
- Wegner, D. M. 1987. Transactive memory: A contemporary analysis of the group mind. *Theories of Group Behavior*. Eds. B. Mullen and G. R. Goethals. Pp. 185–208. Springer.
- Wilson, R. A. 2005. Collective memory, group minds, and the extended mind thesis. *Cognitive Processing* 6(4):227-236.

A theory of collective self-awareness

ANNA MOLTCHANOVA

Carleton College

ABSTRACT

My goal in this paper is to outline and motivate a theory of collective self-awareness. Collective self-awareness is neither supra-individual nor an aggregate of instances of I-awareness. It is comprised of individual loci of we-awareness. A small face-to-face group is collectively self-aware when all group members (1) are we-aware from their first person perspective and (2) include the same participants in the “we” in the events experienced, from their first person perspective, as a “we.”

We-awareness is different from an individual awareness that “I” is part of a group. In the first-person perspective of each group member, we-awareness is an immediate awareness of a “we” that is acting or experiencing, in which the subject is pre-reflectively represented as a “we.” I clarify what “we-aware” means and outline but not discuss in detail why the inevitably diverging representations of experience from different loci of we-awareness do not interfere with collective self-awareness.

I first motivate my theory of collective self-awareness. What appears problematic about claiming that one is experiencing the subject as collective or something is experienced as a “we,” is that group bodies and minds are made up of discrete units, even if connected via certain relations. While individual self-awareness involves proprioception, kinesthetic and self-location, any kind of “we-awareness” would appear to lack these features, because there is no single phenomenological/spatial vantage point of experience for multiple bodies. For each unit it seems impossible to be aware of what other units of the “we” are experiencing in an unmediated way and hence be “we-aware.”

To deal with this difficulty, I discuss the empirical findings that support the existence of the allocentric reference frame, group proprioception as well as shared attentional frame and intentionality. Empirical studies confirm that those experiences to which each member of the group seems to have access only of their own—one’s reference frame, proprioception and the mode of access to what it feels like to have an experience, as well as the neurophysiological traits of an unreflective goal setting and task execution—may change for human individuals in the presence of others.

To illustrate, one unconsciously switches from processing stimuli in an egocentric reference frame to an allocentric one when one detects that another person pays attention to the same object (often to be assessed or manipulated): individuals start operating within a different spatial orientation in their view of the situation and they process information as if looking at the objects from the perspective that does not originate from where their own body is located.

To clarify, one can be I-aware or we-aware in an allocentric reference frame. In a hostile action, like in boxing, one is normally not we-aware but each member of a team trying to control their hurtling down the track in a bobsleigh would most likely have a genuinely collective experience and thus be we-aware. The immediate experience of the subject of boxing and of the bobsleigh team’s ride are different although based on similar neurophysiological under-

pinnings. This is where the phenomenology of the subject from each first-person perspective plays a role. To give another illustration: the ecological content of perception that specifies one's own embodied position in the environment can extend beyond one's body. The bobsleigh team during their run, or demonstrators linking elbows during a march are just two examples of groups that possess ecological we-awareness.

The difference in the representation of the subject in different loci of awareness eliminates collective self-awareness. If A experiences the activity as a genuinely collective subject ("we are doing X"), but B's experience is "I am doing X" (an individual subject), B and A are not collectively self-aware of doing X.

Imagine an ER waiting room in which A thinks that A and B are waiting together to hear the news about the same patient (although A and B are not acquainted). B is indeed waiting to hear the same news, but B thinks A is waiting for something else. For A, the experience is "We are waiting to hear about Q," but for B, it is "I am waiting to hear about Q." A (who thinks *they* are waiting together) is wrong about collective self-awareness. True, there is something it is like to be "I" in "I am waiting," and there is something it is like to be "we" (as experienced from a first-person perspective) in "We are waiting," and A may really feel the latter, but to be aware of one-self, this "one" has to be aware. The "one" of the "we" is actually two: A and B, and both A and B have to be we-aware to be collectively self-aware. Since B is not we-aware, there is no "we" that is collectively self-aware of the "we."

Even if members of the same group are collectively self-aware, simply by position, they would have different representations of the activity. In the last part of the paper, I outline why collective self-awareness, while requiring that the "we" in we-awareness has to be of the same set of individuals for each member, does not require that the content of experience is identical for each of them and sketch how different representation of experience may be reconciled in collective self-awareness.

Collective Intentionality as the common ground of normative experiences

GLORIA MÄHRINGER
Lund University

ABSTRACT

We take certain facts to be reasons. The perception of a reason comes with a psychological pull. We feel driven towards the action or attitude for which we believe to have a reason. In this talk, I will examine the psychology of normative experience and argue that collective intentionality is the basis of any kind of normative experience. While normative experience can occur both in the “I-mode” and in the “we-mode”, it is always based in a reality that only exists in virtue of collective intentionality. The common basis of normative experiences explains why there can be “spill-over” effects between different kinds of normativity, which are often hard to make transparent, and even harder to combat in their detrimental dimensions.

Searle distinguishes between brute facts, existing independently of human intentionality, and social facts, only existing in virtue of collective intentionality. According to Searle, our experience of rights and obligations is due to reason-giving social facts. Such facts endow us with deontic powers. Positive deontic power consists in the right to do something that would otherwise be impossible to do. Negative deontic power consists in the obligation to refrain from something you could otherwise do. Burman argues that deontic facts are not the only type of power-related social facts that are normative. She points out that we should also be aware of telic facts, imposing teleological normativity on us. Examples are facts such as gender or social class. Sometimes, we perceive such facts also as reasons counting in favor of a specific behavior, though we do not recognize them as facts constituting special rights or obligations (at least not in most legal systems of the modern western world). However, their power may have an indirect effect on the deontic powers that an individual has. Thereby, telic powers can indirectly influence actual rights and obligations.

I will argue that, though a systematic distinction between deontic and telic facts is useful to make opaque mechanisms transparent, deontic and teleological normativity share a common ground. Recognizing this common ground helps us to explain and understand the problematic translation of telic power into deontic power and gives us a more critical understanding of normative experience in general. Korsgaard argues that the experience of normativity comes with the existential demand for self-constitution. Self-constitution is an essentially social process, presupposing a relation to others and the ability to perceive reasons as shared. I think that Korsgaard depicts the peculiar phenomenology of normative experience in a plausible way. The self, the subject of willing itself, is at stake when it comes to following the pull of obligation or not. The experience of “obligation” at its developmental origin is existential. An experience of obligation is inherent in first-personal volition itself. The distinction between legal and moral, deontic or telic is only secondary and mediated. However, I think that first-personal volition can occur both in the

singular (I-mode) and in the plural (we-mode). None of them is prior to the other. But collective intentionality is prior to both. The existence of this common ground explains why agreements on the legal/deontic level may fail to affect conditions on the telic level, while expectations on the telic level may well affect processes on other levels.

Understanding the ground of normative experience can further point the direction how to combat the detrimental effects of telic power.

REFERENCES

- Andersson, Åsa (2007): *Power and Social Ontology*. Malmö: Bokbox Publications.
- Baier, Annette (1997): *The Commons of the Mind*. The Paul Carus Lectures Series 19.
- Burman, Åsa: *A Taxonomy of Social Facts* (manuscript)
- Korsgaard, Christine M. (1992): "Creating the kingdom of ends: Reciprocity and responsibility in personal relations." In: *Philosophical Perspectives* 6, pp. 305-332.
- Korsgaard, Christine M. (1993): "The reasons we can share: An attack on the distinction between agent-relative and agent-neutral values." in: *Social Philosophy and Policy* 10 (1), pp. 24-51.
- Korsgaard, Christine M. (1996): *The Sources of Normativity*. Cambridge: Cambridge University Press.
- Korsgaard, Christine (2009): *Self-Constitution. Agency, Identity and Integrity*. Oxford: Oxford University Press.
- Searle, John R. (1995): *The Construction of Social Reality*. London: Penguin Books.
- Searle, John R. (2010): *Making the Social World. The Structure of Human Civilization*. New York: Oxford University Press.
- Skeggs, Beverley (2004): *Class, Self, Culture*. New York: Routledge.
- Tuomela, Raimo (2002): *The Philosophy of Social Practices: A Collective Acceptance View*. Cambridge: Cambridge University Press.

Institutions as Shared Plans: Practical Reasoning in I and We Modes

PEKKA MÄKELÄ, RAUL HAKLI AND S.M. AMADAE

University of Helsinki / University of Helsinki / University of Helsinki

ABSTRACT

Our paper investigates the recent attempt by Francesco Guala and Frank Hindriks (Hindriks & Guala 2015, Guala 2016) to understand institutions by arguing that constitutive rules can be reduced to regulative rules. Their argument relies on accepting standard decision theory consistent with orthodox game theory. Guala argues that the norms which structure institutions are the equilibrium solutions to games and are successful if they solve coordination problems and mixed motive games. He draws on shared intentionality literature, especially John Searle (2010), to argue that the normativity characteristic of institutions need not depend on irreducible collective intentionality. This means that actors perceive that tokens (e.g., dollar bills) represent a specific type (e.g., money), and then act in accordance to this common representational scheme. Searle allows that “there are some forms of collective intentionality which are reducible to I-intentionality plus mutual belief” (Searle 2010, 58). Guala then draws the inference that “this line of argument, if correct, demonstrates that collective intentionality is not necessary for the existence of institutions in general.” We interpret Searle’s point differently and explore another consequence of Guala’s limitation to orthodox game-theoretical framework, namely its lack of the concept of intention. Belief-desire-intention (BDI) architecture encompasses and further allows for more complex action including I-mode (and we-mode) planning that is not possible in decision theory (Pollock 1995, 175-199). Thus we propose to investigate how planning within the framework of BDI architecture introduces a concept of commitment within I-mode reasoning that is absent in decision theory. Commitment can be compared to deontological constraint (Bratman 1987, 166-167). We next discuss how introducing commitment into I-mode practical reasoning has significant implications for understanding we-mode action underlying institutions. We argue that Hindriks and Guala have not fully demonstrated that their reduction works in all cases and we try to clarify some of the points of disagreement between Searle’s and Hindriks & Guala’s accounts of social institutions.

REFERENCES

- Bratman, Michael. 1987. *Intention, Plans and Practical Reasoning*. CLSI Publications.
- Guala, Francesco 2016. *Understanding Institutions: The Science and Philosophy of Living Together*. Princeton University Press.
- Hindriks, Frank & Guala, Francesco 2015. Institutions, rules, and equilibria: A unified theory. *Journal of Institutional Economics* 11, 459-480.
- Pollock, John. 1995. *Cognitive Carpentry: A Blueprint for How to Build a Person*. MIT Press.
- Searle, John R. 2010. *Making the Social World: The Structure of Human Civilization*. Oxford University Press.

Objects, People, and Powers: The Deontic Theory of Money

BENJAMIN NEESER

University of Geneva

ABSTRACT

This paper addresses the freestanding Y-terms objection to John Searle's social ontology (Smith and Searle 2003), focusing on the case of money. It argues, contra both Barry Smith and Searle's response, that a deontic theory of money is the most promising answer to that objection. A cornerstone of John Searle's (1995) theory of social reality is the claim that all institutional objects "bottom out" in some brute basis, i.e. a concrete entity, that serves as its ground, and to which status functions are assigned. This naturalistic, monistic view of institutional reality is expressed in Searle's "X counts as Y in C" formula for constructing institutional objects, where X is a brute entity and Y an assigned status function. Barry Smith's so-called "freestanding Y-terms" objection challenges Searle's naturalism by highlighting a host of cases —property rights, fiat boundaries and most notably electronic money— where institutional objects exist without any brute basis.

The objection has spawned a literature of its own (Thomasson 2003; Hindriks 2013; Smith 2014; Searle 2003, 2006, 2010). However, little attention has been paid to what the aforementioned counterexamples tell us about the nature of the putative institutional objects at hand, especially (electronic) money. This paper comes back to Smith's original argument for counting electronic money as a freestanding Y-term, and uses it to explore the relationships between electronic money, paper money and their alleged brute bases. It argues that the case of electronic money sheds light on how money depends on various institutional objects typically attached to it (bank notes, bank statements, etc.) and ultimately on a misconception of Searle's social ontology. Under the view defended in this paper, money is never a thing counted as such in some context, but a bundle of deontic powers of a specific kind, whose existence maintenance, and change in their distribution over social agents typically require the existence of distinct institutional objects: economic documents.

Smith's argument points out discrepancies in the respective survival conditions of electronic money and any of its putative brute bases. Neither bank statements or ledgers, nor magnetic traces on a bank's hard drive can be the X-term that is counted as money since it would survive their destruction. These should instead be seen as representations of money. In the case of paper money (and material currency more generally) on the other hand, no such discrepancy is to be found, and the usual story goes through: money just is – literally – any concrete thing we are willing to count as such.

The usual morale is that if one wants to stick to the standard account for paper money, and provided that electronic money really is devoid of a brute basis (a claim Searle attempts to resist), one must adopt a form of disjunctivism about money's nature (paper vs. electronic) and existence conditions (with vs. without a brute basis). This is Smith's favoured option. This paper argues that this is both worrisome – a simple, unified view about money should be

preferred – and unnecessary – even paper money needs not be naturalistically accounted for. In other words, money is not a thing, whichever forms it takes.

Instead of trying to find a brute basis for electronic money, this paper contends that ‘paper’ money has none either, although just like ‘electronic’ money, it depends on some regular institutional objects. The argument goes as follows. Identity (or material constitution) is not the only relation that can explain the match in survival conditions between ‘paper’ money and a bank-issued piece of paper it supposedly is (or is constituted by). A weaker relation (entailed by both identity and constitution, but entailing neither) of existential dependence explains just as much, while allowing to keep them separate. What does have a brute basis is the bank note, but bank notes, similarly to bank statements, are documents recording money’s existence, not money itself.

A different dependence relation to documents also explains the mismatch in survival conditions between ‘electronic’ money and bank statements. Smith’s argument does not reveal two kinds of money, but two ways in which it can depend on the material artefacts of institutional reality. ‘Electronic’ money merely depends generically on some bank statement or other (as long as one exists, it does too), while ‘paper’ money depends rigidly on a unique note. This difference in turn is explained by the way each document identifies money’s owner. Direct identification (through a unique designator like a proper name) in bank statements allows for document change or destruction, whereas flexible, descriptive identification (as ‘whomever holds this note’) does not (compare with a passport as opposed to a nameless safe-conduct).

Most importantly, this approach provides room for asking anew what money is and how it comes to existence. On Searle’s theory, we only ever construct institutional reality because of the structure of deontic powers it gives rise to (and in fact, given his broadly naturalistic commitment, deontic powers are the only kind of thing that we can create by collective agreement). Institutional reality is primarily one of people and the deontic powers with which we endow them, and secondarily one of objects whose functions are epistemic (mnemonic, even) and symbolic. They record the rights and powers of certain people, publicly represent them, and contribute to their maintenance and manipulation. Money then, this paper proposes, is a bundle of economic deontic powers (or licenses, obligations), created through collective agreement, and both bank notes and bank statements are mere records on which these powers nonetheless depend.

This last point might seem at odds with the very spirit of the objection this view stems from. But this impression begs the question against Searle’s conventionalism about deontic powers: they only exist insofar as people collectively agree on them, and in the absence of any record (including people’s memories), it seems they do indeed cease to exist. This latter point shows how the proposed approach to money generalizes to other powers involved in Smith’s list of freestanding Y-terms, such as property rights.

Nostra Res Agitur? Programmatic Contexts and The Disentangled We

HENNING NÖRENBERG

University of Rostock

ABSTRACT

In most discussions on collective intentionality, the usual frame of reference is constituted by more or less direct face-to-face-interactions: painting a house, walking together, playing a piano-violin duet, etc. At the same time, it is often supposed that, for the most part, the findings pertaining to those smaller groups can be applied to larger communities such as corporations, political parties, populations, armies, and so on. However, some of those larger communities transcending face-to-face interactions (and referred to in pronouncements such as “We can handle the migration crisis”) seem to feature a peculiar form of we-intentionality that I shall call “disentangled we”.

With regard to that form a couple of questions need to be answered. Given that any concept of collective intentionality that failed to account for the relations between the subject and those others would look somewhat odd, one could ask whether one actually could think of impersonal and anonymous social phenomena as genuine forms of intentionality in the we-mode. What could justify the reference to a we in pronouncements such as “We can handle the migration crisis”? How to conceive such a thing as the disentangled we? To what extent is it parasitic on face-to-face-based we-intentionality which is often regarded as the most fundamental model of the we?

In the following, I provide answers to these questions by introducing the concepts of situation and communal situation as they have been developed in the phenomenological tradition (Husserl, Stein, Walther, Heidegger, Schutz, Schmitz, and others). Basically, the disentangled we is conceivable as a specific type of a communal situation. As I will show, the pivotal point of the disentangled we consists in the subject’s being immersed in a programmatic context shared with others.

In section 1, an example from the end of Peter Strawson’s “Persons” is briefly introduced as a paradigm case of the disentangled we: Arguably, the “we” in “we have taken the citadel” does not only refer to the individuals who actually “did it”, but also to those who at that time were occupied with organizing the supplies of combat boots. In the remainder of the talk, I shall try to make sense of the general implications of Strawson’s example.

Section 2 briefly introduces the concepts of situation and communal situation. Situations provide an ontological model to frame central insights about intentionality from the phenomenological tradition as well as analytic philosophy of mind. They consist of a holistic meaningfulness in which various aspects of meaning (states of affairs, programs and problems) are implicated and can more or less adequately be grasped by the subject. In communal situations the relevant aspects of meaning are not concentrated on the subject only, but apply to others as well.

In section 3, the specific features of the disentangled we as a particular type of communal situations are outlined by generalizing the implications of Strawson’s example. This type is

primarily defined by a contexture of programmatic aspects of which some may be directed to concrete goals whereas others may consist in more goal-independent and 'timeless' practices. Rather than the group members being explicitly jointly committed to each single programmatic aspect, the members' respective cared-for self-images in terms of tacit pretensions and prospects of what they are after or what they wish to avoid are embedded in as well as attuned to the programmatic aspects relevant to the communal situation. (In the context of Strawson's military example, this would certainly have to be addressed as indoctrination, whereas other cases may involve less problematic forms of alignment.) Under these conditions, the sense of we that ought to be in place even when the individual refers to members of her group is mediated by the relevant programmatic context. The other members of our group are not known as unique persons but rather as typified others to whom the relevant programmatic context also applies.

In the conclusion of the talk, I return to the question whether the disentangled we is parasitic to more fundamental and direct forms of we-intentionality. Arguably, communal situations of the type that Strawson has in mind are prepared in and by other communal situations of a less anonymous kind, i.e. situations which precisely entail direct face-to-face interactions. However, the point will be that the individuals involved encounter each other rather in terms of their respective roles prescribed by the programmatic context. In a certain sense, those roles antecede the individuals who fill them, though these roles are of course further specified in terms of how a concrete individual actually fills them. In this sense, the 'we' in terms of such a programmatic context has its own originality as it shapes the various face-to-face encounters embedded in this context.

Social reality without levels

GIANLUCA POZZONI

Universita Degli Studi di Milano

ABSTRACT

This paper seeks to contribute to the debate on the metaphysics of social reality by making the case for a “de-leveilling” of the social. Whether non-individual social objects (e.g. institutions) are granted relative causal autonomy from their components or rather are deemed to be explainable in terms of the aggregate behaviour of individuals, questions about their ontological status are typically framed the a grid that presupposes “levels of the social”. In one variant or another of this idea, individuals are usually taken to be the “building blocks” of social reality which, taken together, constitute “the micro level”; non-individual social objects, on the other hand, are often referred to as “supra-individual” entities that lie at “the macro level” of social reality. Accordingly, the conundrum for much ontological debate is how to account for the ontological persistence and apparent causal efficaciousness of social entities while at the same time maintaining that they are fully made up of individual entities.

Recently, Epstein (2015) has identified “the consensus view” of social theory with the universal assumption of “ontological individualism” while allowing for debate on “explanatory individualism”, i.e. the idea that all social phenomena should be explained solely in individualistic terms. According to ontological individualism, social facts are exhaustively determined by facts about individuals and their interactions, and the consensus view assumes this thesis as a trivial one. Epstein identifies the main alternative to the consensus view in what one author (Guala 2007) has termed the “Standard Model of Social Ontology” (SMOSO), which encompasses theories otherwise as diverse as Ian Hacking’s (1999) and John Searle’s (1995). SMOSO’s account of the relationship between individuals and social facts, in particular, is based on the premise of performativity. On this assumption, individuals hold beliefs about social facts (e.g. “bills issued by the Bureau of Engraving and Printing count as money in the United States”, as per Searle’s famous example) and are collectively capable to enact those beliefs through certain uses of language, thereby putting those very social facts into existence.

As one can see, however, SMOSO’s theory of performativity is a thesis about social facts, not entities. Thanks to performativity, social facts do acquire causal autonomy from the individuals who enact them (e.g. the purchase power of paper money), yet they do not involve newly generated social objects in their own right. As such, SMOSO is tied to an individualistic thesis about the entities that populate the social world. Not unlike the consensus view, SMOSO separates the question about the nature of social entities from issues about the causal explanation of social facts. Like in the consensus view, the causal reality of non-individual facts does not depend on the existence of social objects.

If there is nothing to social reality other than individuals, however, it is not entirely clear why social facts should not be entirely reducible to facts about individuals. Indeed, it is a basic premise of realism about causation and a well-established scientific practice that explanations should be formulated in terms of the entities that produce those facts. As a consequence, one should be able to reformulate facts about the purchase power of money in all their occurrences

with facts about the collective intentionality of the individuals that perform their beliefs about money. (And, one may add, facts about the Bureau of Printing and Engraving with facts about the collective intentionality of individuals that perform their beliefs about the authority to issue paper money, and so on.)

For this reason, some have advocated the idea that the causal autonomy of social facts can be ascribed to the causally generative properties of non-individual entities. To underpin the idea that non-individual entities are endowed with generative properties, it is customary to refer to the notion of “emergence” and defend the idea that social aggregates possess “emergent” properties that are not possessed by any of their components (for a recent version of this position, see Elder-Vass 2010). The very idea of emergence, however, presupposes that non-individual entities are of a “macro” kind, and while their properties are not the mere aggregate of the properties of “micro” entities (individuals), they still result from them. In principle, therefore, one should be able to explain them in terms of facts about “micro” entities.

This paper will argue that this riddle can be solved by doing away with the idea of “levels” of social reality. Understood in realist terms, the causal reality of social facts does in fact presuppose that generative properties can be ascribed to social entities. Moving from this assumption alone, it is possible to adopt a deflationary approach to social ontology and limit the question about the ontological status of social objects to their causal properties. To the extent that social objects are causally generative, it is argued, they can be granted existence on a par with individual entities. This move rests on an assumption that is somewhat close to what is sometimes referred to as “the Eleatic principle” or “Alexander’s dictum”, namely, the principle that “to be real is to have causal powers”. Yet it does also accommodate some assumptions that serve as a justification for the idea of levels of the social, namely, that social objects exist at different scales. Yet such multi-scaled social reality concerns the “magnitude” and internal structure of social objects, not the causal interaction between them.

REFERENCES

- Elder-Vass, Dave. 2010. *The Causal Powers of Social Structures: Emergence, Structure, and Agency*. Cambridge, UK: Cambridge University Press.
- Epstein, Brian. 2015. *The Ant Trap: Rebuilding the Foundations of the Social Sciences*. Oxford: Oxford University Press.
- Guala, Francesco. 2007. “The Philosophy of Social Science: Metaphysical and Empirical.” *Philosophy Compass* 2 (6): 954–980.
- Hacking, Ian. 1999. *The Social Construction of What?* Cambridge, MA: Harvard University Press.
- Searle, John. 1995. *The Construction of Social Reality*. New York: The Free Press.

Macro phenomena without macro manifestations – why the multiple realization argument is incompatible with social science

GUSTAV RAMSTRÖM

Stockholm University

ABSTRACT

My paper argues that the multiple realization argument cannot be applied to social science (as suggested by for example Kincaid, Sawyer and List and Spiekermann). I argue that whilst the argument is potentially applicable to physical sciences which deal with macro manifest macro properties (type 1) it is not applicable to the social sciences where we deal with micro-based macro properties (type 2).

That is, a property of X which we define as a “macro property” may refer to X taken as a unified whole (type 1) or of X taken as a composite system (type 2). Whilst type 1 macro properties refer to properties which emerge at a higher level of manifestation given lower level circumstances type 2 macro properties label, summarize and describe the lower level circumstances which characterize the disposition and structure of the system. If we, for example, identify a diamond as “hard” this would be a type one macro property whilst if we characterized the crystalized structure which the carbon atoms of a diamond constitute as “concentrated” this would be a type two macro property. Both types are macro properties in the sense that they are ascribed to the system as a whole rather than to any of its parts in particular. However, they also display significant differences in that type 1 macro properties are identified at a distinct “macro level of manifestation” (or macro level of zoom/grain) whilst type 2 macro properties are inferred from circumstances in the micro-base (i.e. the properties, activities and interactions of micro elements).

In many scientific disciplines micro investigation and the knowledge about the micro base (the properties, behavior and interrelations of micro entities) acquired through such investigation is used to explain or understand macro manifest properties, or if you will, “higher level” properties. A diamond is thus hard (type 1) because it has a concentrated atomic structure (type 2) and the boiling point of water (type 1) depends on the behavior and interrelations of water molecules (type 2) etc. However, this is not the case in all scientific disciplines. In social science, for instance, micro-structural properties are not used to explain higher level phenomena, indeed the types of phenomena and properties we are interested in as social scientists do not even have macro manifestations.

Let’s use the often invoked (but ill understood) relationship between social science and philosophy of mind to illustrate this lack of macro manifestations in social science. Compare the relationship between psychological states (macro) and neural activity (micro) in philosophy of mind/psychology to societal phenomena (macro) and the activity of individuals (micro) in social science. Whilst a mental experience such as “pain” refers to a macro psychological state rather than to a certain type of micro neural activity, there is no equivalent macro

manifestation of a social concept like “riot”. That is, no higher level events than the activities of human beings can be said to constitute an empirical manifestation of a “riot” whereas the feeling “pain” is manifested both at the macro level (in terms of a psychological experience) and at the micro level (in terms of neural activity). Unlike a “riot” we can thus identify pain at the macro level after which we can start investigating how it is manifested or realized at the micro (neurological level). Whilst “pain” can be identified without resorting to neurological level observation/information, a “riot” could not be identified from a vantage point where we didn’t have information about the activities of individuals.

The same is true for all social entities, properties and phenomena. That is, no families, companies or nations manifest themselves in any other form than in terms of the activities, interrelations and mental properties etc. of individuals in physical settings, no matter how far we zoom out. We could, for the same reason, never identify a society as democratic, a company as hierarchical, a nation as torn by civil war or a family as dysfunctional etc. without information about the activities, interrelations and properties of individuals. In social science we thus never leave the micro-base. The level of acting, interacting and experiencing individuals is the only (or rather highest) empirical level we ever deal with. Higher levels of analysis, is thus just that, higher levels of analysis, not higher levels of manifestation or higher levels of empirical reality.

I argue that only macro manifest phenomena, i.e. phenomena which have higher level manifestations, can be found to have multiple and unpredictable micro level realizations (i.e. be “wildly disjunctive”). A micro-based macro phenomenon (e.g. a social phenomenon) is necessarily inferred from micro level states and events and its realization must thus, explicitly or implicitly, be determined before it can be identified. The realization of a variable in empirical social science is thus the opposite of “wildly disjunctive”, it has a finite set of realizations which are determined a priori. “Multiple realization” as this concept traditionally has been understood in philosophy of science, i.e. as an empirical circumstance characterizing the relationship between higher and lower level phenomena/properties does not exist in social science. It can only be applied to micro-macro relations where micro-circumstances actually give rise to higher level phenomena. Indeed, the transference of the multiple realization argument from philosophy of mind to philosophy of social science is indicative of a conflation between macro manifest and micro-based macro properties. That is, those who apply the multiple realization argument to social science treat “levels of analysis” (e.g. social vs. individual level facts) which are inferred from/refer to the same manifest (or “empirical” base) as if they were “levels of manifestation” (e.g. facts about molecules vs. facts about atoms), i.e. distinguishable empirical levels of reality.

Collective Agency and Individual Criminal Responsibility

EIRIK JULIUS RISBERG

Falstad Centre

ABSTRACT

1. From individual criminal responsibility to crimes against humanity?

We generally accept that political leaders and high-ranking officers who are responsible for the political and strategic decisions in resorting to war also bear criminal responsibility for the terror and mayhem they thereby unleash. But what about individual, lower-ranking soldiers, policemen and prison-guards who are enlisted in a system with a built-in requirement of strict obedience to those upwards in the chain of command? The question is far from new, and received its perhaps most poignant formulation in the Nuremberg trials immediately after the Second World War – with the most contested discussion arising from Hannah Arendt’s portrayal of Adolf Eichmann in a series of articles for *The New York Times*. But while lower-ranking soldiers and prison guards may be tried and convicted of crimes such as torture, murder and rape, how can they be convicted of crimes against humanity, crimes that are by definition widespread and systematic and thus collective? Or as Larry May puts the question: “what would need to be shown in order for a soldier to be criminally liable for violating the rules of war and international humanitarian law” (May 2007: 235). Taking as its starting point a verdict handed down by the International Criminal Tribunal for the former Yugoslavia (ICTY) against five lower-ranking individuals, this paper seeks to address how we get from individual criminal responsibility for human rights violations committed in war to a moral argument that conceives of these crimes as crimes against humanity. In doing so, it raises the question whether collective action needs to postulate the existence of group agents as subjects of shared intentions, or whether we can do with the notion of a we-intention. If a notion of we-intention is sufficient to explain collective action, what requirements or constraints are placed upon the notion of a we-intention that will allow us to deduce a conception of individual criminal responsibility from it?

2. The case of Omarska prison camp

In 2001, four prison guards and one civilian at the notorious Omarska prison camp in Bosnia were convicted of crimes against humanity for extensive and systematic torture of prisoners in the camp during the Bosnian war between 1992 and 1995 (Simons 2001). According to the ICTY, the five had committed atrocities against inmates that clearly qualified as torture and were thus criminally responsible for serious violations of domestic criminal law. The definition of torture that the ICTY set down was that it was a type of action that intentionally sought to inflict pain on detainees for prohibited purposes, e.g. the obtaining of information or a confession, or simply to humiliate (cf. May 2007: 239). The legal problem facing the prosecution was to make the case that the torture that took place at Omarska was an international war crime, and to distinguish it from torture *per se*. The solution adopted by the court, was to argue that for torture to be considered a war crime, it “must be related to some purpose of

war as part of the special intent that linked the acts of the guards to larger prohibited goals" (May 2007: 239). At the same time, however, the court argued that it was not necessary in order to be held criminally responsible for war crimes that the individuals responsible for the torture were themselves aware of or conceived of the torture in terms of the overall war-aims. The individuals responsible for the torture did not need to share in or even have knowledge of these war aims, whose purpose the camp and the torture was supposed to further, as long as they willingly participated in a system or collective that intentionally used torture as part of its war efforts. But if the individuals responsible for the torture intended something else in torturing prisoners at Omarska than what the collective they were part of did, how can it be argued that they share in the collective guilt of the system?

3. Shared intentionality and collective responsibility

The immediate problem then, is to explain how individuals can be said to share in the collective moral responsibility for the atrocities committed understood as intended to further the war-effort, that is, as crimes against humanity, when they themselves did not conceive of their actions in that way. Our first task will thus be to understand how we construe collective intentionality, and secondly, whether the type of concerted action by the prison-guards is a type of collective intentionality? Margaret Gilbert has argued that a "joint commitment ... is not a conjunction of a personal commitment of one and separate commitments of the others. Rather, it is the commitment of them all" (Gilbert 2013: 64). And further, that this makes the parties "jointly committed to X as a body" (Gilbert 2013: 64). Marija Jankovic and Kirk Ludwig, on the other hand, make the case that we "don't need group agents for collective action, and we don't need them as subject of shared intention" (2016: 3, preprint). Instead they argue, we can reduce the problem of shared intention to that of we-intentions (cf. Jankovic and Ludwig 2016: 3, preprint). Both views face some difficulties in explaining individual criminal responsibility for war crimes, however. Gilbert's view seems, on the face of it, to require a stronger commitment to the existence of a collective over and above the individuals comprising the body, but it has the advantage that it does not require the parties to hold any individual intentions. Jankovic and Ludwig's view does away with shared intentionality in favour of we-intentions, but explains the latter as individuals intending to X. But can either of them explain the shared intentionality of the prison-guards at Omarska in a way that makes them individually criminally responsible?

The Practical Significance of Group Reasons for Individual Agents

ABE ROTH

Ohio State University

ABSTRACT

When is a reason that we have relevant for what I (should) do? It seems obvious that if I am one of the we, and we have a reason to ϕ , then I have a reason to do my part. But matters are more puzzling than they appear. Consider:

Case 1: Making a difference (or not). I rode my bicycle to work today. More often than not I enjoy it, and that's a reason for riding. Moreover, it contributes to my physical and psychological wellbeing. So those are additional reasons. A further reason - one that is important for what concerns us here - has to do with the environmental damage associated with a car commute. How does this translate into a reason for me to ride my bike? Whether I ride my bicycle or take a car makes no significant difference in the large-scale environmental matters of concern to me, such as that of global warming. My contribution is a drop in a pool of Olympic proportions. How can the environmental considerations really serve as a reason for biking to work in the face of this skeptical "makes no difference" argument? We collectively, as a group, have reason to ride our bikes; it makes a substantive difference if many or all of us ride our bikes instead of commute by car. But since it makes no real difference whether I join in, it's not clear how this fact about the group's reason for action translates into a reason for me.

Case 2: Hi-Lo. Consider a scenario familiar to those acquainted with the literature on team reasoning. You and I each have two available strategies for interaction, A and B. If we both coordinate on A, then it's the best for all concerned. If we both do B, that's second best, but much worse than A. And if we fail to coordinate and one of us does A while the other does B, the result is quite bad. In this scenario we should rationally coordinate on each of us A-ing. That's what we have most reason to do. But how does this fact about our reason translate into what I am supposed to do? Should I implement A? Not if you choose B; in that case, I should choose B. Suppose that what you do will be determined by what you think we should do (you are what the literature refers to as a team reasoner). Then you would pick A. So, if I know that you're a team reasoner, then it's clear that I should opt for A. But notice that in this case the group reason for choosing A doesn't do any work as a reason for me to act. That's because if I know that you're a team reasoner, ordinary individual instrumental reasoning will lead me to opt for A, for that is what maximizes value given my expectations. Thus, it seems that the only way in which a group reason would have any significance for how I should act is one where the group reason makes no difference.

Both cases illustrate group reasons that appear to have little practical significance for individual agents. Together they paint a rather skeptical picture of group reasons. Against this, I will first argue, regarding Case 1, that one can sometimes act for a reason that is a group's reason. That is, I will argue that a participant in a group is sometimes entitled to reasons that are otherwise inaccessible to them. For example, a reason for riding one's bike is grounded

in environmental considerations. One does not by oneself have such a reason, at least not on any single occasion. Only groups that are large enough have such a reason. But an individual can, nevertheless, be entitled to this reason, and this is what enables her to act for this reason. But when is one entitled to the reasons of a group? To get a start in answering this question, I consider diachronic cases of entitlement within an individual, and interpersonal entitlement to reasons in shared agency.

As for Case 2, some might suggest that the team reasoning mentioned there offers individuals a way of gaining access to group reasons that they otherwise would not have. My presentation was meant to raise a challenge for this approach: the problem I face in predicting whether you will opt for A becomes the problem of assessing whether you are a team reasoner. And if I have evidence that you are, then I have no need for the group's reason, nor for team reasoning: I would already have an individual reason for opting for A. How can we address this challenge to team reasoning and the significance of group reasons in this scenario? The answer, I contend, is that my warrant for thinking that you are a team reasoner is non-evidential. That is, it is a form of epistemic entitlement that comes with one's exercise of a distinctive form of agency when acting with others; as such, this entitlement is not available to an individual agent who is acting on her own. This might strike one as a cheat. But I argue that there is precedent for this. Even in the case of individual agency, there will be instances of non-evidential warrant for belief. I illustrate the point with a discussion of mundane decisions that issue in future directed intentions.

By invoking a distinctive form of agency exercised when acting with others, we demonstrate how group reasons can have a significance for agents they otherwise would not have. I consider implications for reductionist accounts of group agency in terms of individual agency. Finally, some have suggested that team reasoning offers a conception of joint intention that can be of use for accounts of group agency. I argue, however, that the order of dependence is if anything quite the opposite: as a theory of rationality, team reasoning relies on a distinctive form of joint agency - the one exercised when one acts with others.

Discrimination and Collaboration in Social Networks

HANNAH RUBIN AND CAILIN O'CONNOR

University of Groningen / UC Irvine

ABSTRACT

We take an in-depth look at the dynamics of discrimination in social networks. Using formal modeling, we show that discrimination can arise even without any individual exhibiting any bias. Further, this discrimination can lead members of different social groups to mostly collaborate with in-group members, decreasing the effective diversity of the social network. While this analysis applies to a variety of contexts where collaborations occur in communities whose members belong to different social identity groups, we focus on the example of collaborations in epistemic communities, such as communities of scientists. Philosophers of science have pointed out that personal diversity, i.e., diversity with respect to personal identity markers such as gender, race, and cultural origin, is important for epistemic success (e.g. Longino, 1990).

In part, our analysis is inspired by two empirical results. First, in epistemic communities, certain minority groups – such as women and people of color – may get less credit for joint work (West et al., 2013). Second, minority group members are less likely to collaborate with member of the majority social group and often cluster in sub-disciplines (Botts et al., 2014). Part of our question is: are these sets of results related? Does inequity in academic collaboration lead members of certain groups to self-segregate and thus decrease the effective diversity of collaborative teams? It is notoriously difficult to generate empirical data testing cultural evolutionary pathways. To explore these questions, we instead employ game theoretic models. Such models start with a game, or a simplified representation of a strategic interaction.

To represent division of labor and credit between academic collaborators, we use the Nash demand game. This game involves two agents who divide a resource by each demanding some portion of it. We argue that this is a good representation of academic collaboration. Because academic collaboration involves joint action which increases productivity, it creates a resource of surplus credit compared to solo work (Bruner and O'Connor, 2015). However, this joint action necessitates two types of bargaining. First, actors must decide who will do how much work on the project. Second, actors must determine author order as a proxy for credit. The demands in the game, then, are best understood as requests for author position relative to the amount of work done. An actor who, for example, does the lion's share of the work and requests first authorship makes a fair demand. One who does more work, and requests second authorship makes a low demand, demanding less of the resource, while one who does less work and requests first authorship makes a high demand, demanding more of the resource. If the demands are compatible, each agent gets what she requested. If the demands exceed the total resource (e.g. if two agents make high demands), collaboration breaks down on the assumption that they cannot peacefully agree on a division.

Suppose we have a population with two social groups – women and men, for example, or black and white people. Suppose further that actors can condition their choice of strategy

on the group membership of an interactive partner. In a cultural evolutionary scenario, this induces a situation where separate norms can emerge within and between groups. We follow authors like Young (1993) in labeling emergent patterns of group level behavior as 'norms', though this is obviously a thin representation of real world norms. Within groups, agents will most often learn to all make fair demands of their in-group members. One of three things will happen between groups. Either the groups will come to make fair demands from each other, or else one group will learn to always make high demands and the other to always make low demands when meeting out-group members. We take these two latter sorts of outcomes to represent discriminatory norms as actors treat in- and out-group members differently, to the detriment of one group.

We look at the emergence of these discriminatory norms, particularly the norm where the members of the majority group discriminate against members of the minority group. We start by looking at the emergence of discriminatory norms in fixed collaboration networks, networks where the collaborations each agent has are determined at the start and do not change over time. Such norms commonly evolve, and minority status alone can make it more likely for a social group to be disadvantaged by bargaining norms. Next, we explore the endogenous emergence of collaboration networks in a population that already has discriminatory norms, finding that such networks tend to become segregated to the point where there are no collaborations across groups. Lastly, we examine the simultaneous co-evolution of discrimination and collaboration, where we see partially segregated networks arise, with some members of the majority group discriminating against the minority group.

Our results suggest a connection between evidence that minority group members receive less credit in collaborations and evidence that they tend to collaborate within their own social identity group. Our results also suggest a process by which academic communities will spontaneously un-diversify in the face of discriminatory bargaining norms. Furthermore, they suggest that such norms can spontaneously emerge in academic communities under many conditions, and are more likely to impact minority groups. To ensure the diversity of epistemic communities and collaborations may take concerted effort to fight the social dynamical forces that divide social groups.

REFERENCES

- Botts, Tina Fernandes, Liam Kofi Bright, Myisha Cherry, Guntur Mallarangeng, and Quayshawn Spencer (2014). "What Is the State of Blacks in Philosophy?." *Critical Philosophy of Race*, 2 (2), 224-242.
- Bruner, Justin P and Cailin O'Connor (2015). "Power, Bargaining, and Collaboration." *Scientific Collaboration and Collective Knowledge*. Ed. Conor Mayo-Wilson Thomas Boyer and Michael Weisberg. Oxford University Press.
- Longino, Helen E (1990). *Science as social knowledge: Values and objectivity in scientific inquiry*. Princeton University Press.
- West, Jevin D, Jennifer Jacquet, Molly M King, Shelley J Correll, and Carl T Bergstrom (2013). "The role of gender in scholarly authorship." *PloS one*, 8 (7), e66212.
- Young, H Peyton (1993). "The evolution of conventions." *Econometrica*, 57-84.

Money and Deontic Power

JOSHUA RUST

Stetson University

ABSTRACT

Francesco Guala aptly describes the example of money as an “undisputed classic in social ontology” (Guala 2009, 259). I raise a consideration that speaks against the claim that money is an institutional fact. While I ultimately conclude that money is an institutional fact, seeing this requires going beyond the textbook model of money, as relied upon by many social ontologists.

Money has been described as a numeraire, a medium of exchange, a store of value, a means of payment, a unit of account, a measure of wealth, etc. Of these properties, macroeconomic textbooks typically present the commodity or metallist theory of money, which stresses money’s capacity to serve as a medium of exchange. Objects with an independent use-value that are sufficiently durable, portable, dividable, etc. can be used as a medium of exchange, and, so, a unit of account. John Searle tentatively accepts the commodity theory’s account of money’s nature and origins: commodity money, such as gold, gave rise to contract money, in which the bearers of a given certificate can expect to exchange that certificate for a deposit of gold. These certificates, then, began to be traded in the marketplace. Finally, contract money was replaced by fiat money (Searle 1995, 41-43).

If every institutional fact carries deontological powers, as Searle maintains (Searle 2010, 23-24), then it seems that the gold or silver coins (commodity money) used in primitive exchange scenarios are not institutional facts. A trader has a reason to exchange surplus goods for money things, but the reason is a desire-dependent one--the trader expects to be able to exchange those metal coins or shells for a desired good at a later date, thereby solving the “double coincidence of needs” problem. Neither of the parties, in this primitive exchange scenario, are deontologically bound. Contract money and fiat money, by contrast, carry a deontology. However, the fact that only contract and fiat money carry a deontology implies an unwelcome distinction between the various kinds of money described within the commodity theory: only contract and fiat money are institutional facts.

Frank Hindriks offers a novel solution to this problem, which arises, he contends, because Searle’s notion of deontology is “too restrictive” (Hindriks 2008, 132) and should include second-order Hohfeldian incidents, in addition to first-order rights and obligations. He claims that money provides one with purchasing power, which “is a matter of being able to change the distribution of property rights” (Hindriks 2008, 132).

Against Hindriks, I argue that if a money-bearer is able to change the distribution of property rights, it is because money has persuasive power, not deontological power. The bearer of money is not entitled to another’s property, even if trade results in a realignment of property rights. Still, Hindriks is right to have us focus on second-order powers, and especially the power to change the distribution of property rights. However, the relevant power appears to be, not purchasing power, but the power of alienation—the power to extinguish the money bearer’s own interest. This is a deontological power money has in virtue of the fact that it is a member of the social kind, property.

Every institutional fact implies a deontology and this solution finds money's deontology in the power of alienation. However, the solution raises a new difficulty with the suggestion that money is an institutional fact. On the plausible assumption that the deontology carried by social kinds must individuate that kind, if money entails the power of alienation, and the power of alienation is entailed by property in general, then money is not sufficiently individuated from (other kinds of) property.

I appeal to the credit theory of money to argue that money does carry a unique deontological power. Economist Charles Goodhart describes the credit, cartalist, or claim theory of money, in contrast to the commodity theory of money. The credit theory holds that an object comes to constitute a money thing in virtue of its relation to a sovereign authority. What matters, here, is not that a given coin has an independent use-value, but only that it can be used to pay taxes and other state-sponsored obligations.

So what, exactly, is money on the credit theory? Money is a kind of credit or entitlement--to have money is simply to have the right or permission to eliminate one's tax burden. In this way, money things are institutional facts which are sufficiently individuated from other kinds of property and credit. While my focus is on the status of money things in primitive exchange scenarios, a consequence of this view, accepted by credit theorists, is that private media of exchange, such as Conch Dollars or Bitcoins, are not, strictly speaking, money (See Wray 2015, 111-12). But they could be if they came to satisfy one's state-sponsored obligations.

I conclude by comparing the proposal with that of Guala's, who rightly observes that transferable tax credits make an ideal medium of exchange: "fiat money is sustained by two mechanisms: (1) individuals know that fiat money reduces their transaction costs, and (2) they know that they will be punished if they do not use it in some transactions (e.g., tax payments)" (Guala 2016, 40). The problem with Guala's view is that he treats the state imposition of an obligation on its subjects as incidental to something's being an expression of money--the state's demand for certain tokens happens to prompt acceptance, but is not strictly speaking required for a token to constitute a money thing. On the proposed view, a money thing must necessarily imply a tax credit, even if a primary function of those money things is to serve as a medium of exchange. Otherwise, there do not appear to be any unique deontological markers that can distinguish money from other kinds of property.

From Coordination to Cooperation. Switching to Team Reasoning

JULES SALOMONE

CUNY - The Graduate Center

ABSTRACT

In this paper, I argue that when facing coordination problems, agents have a pro tanto reason to cooperate, where this pro tanto reason derives from conditions constitutive of their agency. Before I can even defend this claim, I must spell out what cooperation takes. I first show that, although insightful, Bratman's account of shared cooperative activity fails to exclude ostensibly non-cooperative Nash Equilibria (I here rehearse and expand on a claim defended by Gold and Sugden). To do so, I construct an example of a (hawk, dove) strategy profile that meets all of Bratman's conditions (namely, and roughly, that all participants are committed to the meshing of their subplans, that they all intend the efficacy of each other's intentions, that no participants is coerced into forming intentions relevant to their shared activity, that a condition of minimal mutual support obtains, and that they keep an eye on each other's actions).

I then proceed to my analysis of the concept of cooperation. I first argue for a series of relatively uncontroversial conditions. On my view, some activity ϕ is cooperative only if: parties to ϕ commit themselves to the meshing of their plans which, taken together, constitute ϕ ; participants are not coerced into committing themselves to the meshing of their plans, nor into forming their plans, nor into acting on their plans; a condition of minimal mutual support obtains; and parties to ϕ are responsive in action when in the process of ϕ -ing. Following Roth, I analyze the commitment to the meshing of plans along these lines. Either I take your desire/intention to act on some plan as settling the issue for us, in which case your wish silences all the considerations favoring or telling against this plan of yours that I might have antecedently contemplated. Or your wish enters in my deliberation as one considerations among others, in which case my commitment to the meshing of our plans dictates that I adjust the weight of your wish in a way that it prevails in my deliberation, or that we try to resolve outstanding conflicts in a manner that is acceptable to both of us.

I then turn to a further necessary condition for cooperation. I argue that, in order to avoid counterexamples to which Bratman's account falls prey, a cooperative activity must be favored by reasons acceptable to all parties from their shared perspective. I first show that this condition leaves open the possibility that two parties' respective commitments to the meshing of their plans are favored by different reasons: on my view, cooperation does not require a perfect alignment of reasons. I then show how this condition keeps at bay non-cooperative strategy profiles. When framing the decision problem from their shared perspective, participants share (defeasible) rational control over each other's actions. As a result, shared perspectives transform any strategy profile (and not just Nash Equilibria) into feasible profiles, provided that they are favored by reasons acceptable to all. As non-cooperative profiles are not supported by such reasons, the condition I argue for excludes them from the extension of my concept of cooperative activities. This line of reasoning de facto commits me to the claim that some version of

team reasoning stands as a necessary condition for cooperation.

Before concluding my analysis of the concept of cooperation, I draw attention to the fact that the adoption of a shared perspective may not by itself lead to the commitment to the meshing of plans. First, individuals reasoning from this perspective, even when satisfying standards of reasonableness, may fail to accept each other's reasons: even agents disposed to cooperation and thereby ready to reason from their shared perspective may soon realize that their plans insuperably conflict. Second, agents reasoning towards reasons acceptable to all may conclude that such reasons favor not committing themselves to the meshing of plans – as it is the case (e.g.) in defenses of free market.

Once the concept of cooperation is on the table, I refine the view that I can reasonably hope to defend under the assumption that 'ought implies can': agents engaged in coordinated activities have a pro tanto reason to team reason. I then stipulatively define coordination as – roughly – situations in which reasonable agents' deliberations include considerations about others' decisions and actions. A consequence of coordination is that the success of agents' actions depends on others' actions.

I finally argue that if control is a constitutive condition of agency – a relatively uncontroversial claim –, then agents ought to cooperate in order to exert joint control on their endeavors. Indeed, as parties to coordinated activities do not share the same phenomenological unity, they cannot use any of the controlling abilities afforded by the possession of a unified consciousness. As a result, in order to live up to the control condition partly constitutive of agency, such parties seemingly have two options. They may either coerce others so as to make sure that the latter's actions contribute to the success of the former's endeavor. Or they may reason towards considerations acceptable to all. I then show that coercion is less of an option than team reasoning.

I conclude with the following observations. First, my argument for cooperation through team reasoning differs from arguments to the effect that prudential reasons only favor cooperative activities. The latter presuppose standards of evaluation of the acceptability of reasons favoring cooperation: only reasons bearing on the well-being of parties are worthy of consideration. By contrast, my argument carries no such presupposition. Second, my argument suggests that uncooperative parties to coordinated activities not only let themselves down (as Grice noticed), but also the other parties with whom they fail to cooperate: in failing to adopt a shared perspective, they prevent the exercise of shared rational control over coordinated activities, and thereby prevent others as much as themselves from behaving as full-blown agents.

Non-observational knowledge and Joint action

GLENDALUCILA SATNE

University of Wollongong

ABSTRACT

It is a commonly held view, both in current Philosophy of Mind and Developmental Psychology, that understanding the nature and development of human knowledge of other minds requires understanding the relationship between such knowledge and characteristic forms of human interaction and shared agency (see recent work by Reddy, Avramides, Gallagher, Butterfill, among others). In this vein, Butterfill 2012, for example, argues that interactive agents have a route to knowledge to other people's mental states that passive observers are not in the position to have. Nevertheless, this very insight can be substantiated in different ways, depending on how one understands the relevant forms of interaction and shared agency.

In this presentation, I am interested in pursuing this idea- that interactants in joint action have a special route to knowledge to other people's mental states- by following a key idea to be found in Anscombe's work on intentional action. Namely, that one typically knows what one is intentionally doing without evidence or observation (Anscombe, Hampshire, Moran). According to this view, you don't know what you are doing by gathering evidence about what you are up to. For example, you don't find out what it is that you are doing by watching yourself do it. Observation comes into it, but rather as an aid (Anscombe 1957, 53). Following this idea, perception is rather an enabling condition in pursuing the action you set yourself to pursue and not the source of the knowledge you have of what you are doing. By contrast, the kind of knowledge at issue is classified by Anscombe as belonging to the realm of Practical Knowledge rather than Theoretical or Observational Knowledge in that it is made intelligible in terms of the agent's possession of reasons for performing a particular action, reasons that provide a point or purpose to the action in question.

Moving to the plural case, this presentation argues that the distinction between observational and non-observational knowledge provides a criterion for distinguishing parallel action from joint action in terms of the kind of knowledge that agents are in the position to have about what others are doing when acting together with them, but do not have when acting in parallel. Consider the following situation: you arrive at the train station in a rush just to learn that your train is delayed and see a few people standing in the quay anxiously checking their watches. You may well learn from observation that they are waiting for the 8.20 train to Sydney. Given your first-personal knowledge – I am waiting for the 8.20 train to Sydney – you may well conclude 'we are waiting for the 8.20 train to Sydney'. Contrast with the following case. You and Hanna are playing tennis on a Saturday morning and when asked what you are doing you answer 'We are playing tennis'. This latter judgement does not seem to be based on your prior, independent knowledge that she, and that you, are playing tennis. Intuitively, such a dual basis could not deliver the kind of knowledge you seem to have — namely, knowledge that you and her are playing tennis together. A tempting hypothesis is then the

following: your knowledge in the second (but not in the first) case is an example of 'practical knowledge' by which I mean that (i) it is not based on observation; (b) it is knowledge whose expressions are expressions of intentions and (iii) its possession is intelligible in the light of the practical reasoning informing the corresponding intention. In particular, it is a case of 'Joint Practical Knowledge'. While some authors have advanced related thesis about the nature of the knowledge involved in joint practical activity (see Laurence 2011, Schmid 2015, 2016) this alternative remains relatively underexplored in the current literature. After making a case for the plausibility of this hypothesis and discussing the differences between the proposed approach and those due to Laurence 2011 and Schmid 2016, the potential of this thesis for understanding and advancing current debates about self-knowledge and other-knowledge in philosophy and developmental psychology is briefly assessed.

REFERENCES

- Anscombe. G.E.M. 1957 *Intention*, Oxford: Basil Blackwell.
- Butterfill, S. 2012 "Interactive Mindreaders". In *Philosophical Studies* 165(3), pp. 841-863.
- Laurence, B. 2011. "An Anscombean Approach to Collective Action". In *Anscombe's Intention*, (Harvard University Press, 2010), edited by Jennifer Hornsby, Frederick Stoutland and Anton Ford.
- Schmid, H.B. 2014. "Plural self-awareness". *Phenomenology and the Cognitive Sciences* 13 (1):7-24.
- Schmid, H.B. 2016. "On Knowing What We're Doing Together. Groundless Group Self-Knowledge and Plural Self-Blindness". In *The Epistemic Life of Groups: Essays in the Epistemology of Collectives* (Oxford: Oxford UP), Edited by Michael S. Brady and Miranda Fricker.

Collective Responsibilities of Random Collections: Plural Self-Awareness Among Strangers

HANS BERNHARD SCHMID

University of Vienna

ABSTRACT

Among the things people should do are things people can only do jointly, together with others. A well-organized society usually has corporate agencies for these tasks. Since crime intervention in larger communities is too big a task for a single Sheriff, we have our Police Departments. The Ambulance is in charge of medical emergencies, for poverty and depravation there are the Social Services, for moral catastrophes in remote corners of the world we have Foreign Aid, and so on. These and similar organizations are the agents to take care of those morally sensitive tasks which no one can perform alone.

We hold these organizations responsible if they fail. In recent philosophical research, it has come to be widely accepted that we are justified in doing so (cf. List/Pettit 2011). One of the arguments for collective or corporate responsibility is that the responsibility in question often cannot be straightforwardly reduced to the responsibility of the individual members of the organization. Though there are usually individual members to blame, blaming them for failing to do their job is different from blaming them for the act or omission in question. The Chief of the Fire Department is responsible for failing to re-organize his department, but you may not simply blame the boss for not putting out the fire in the way you would blame an individual bystander who fails to act with all the necessary means to intervene at her hands. Thus it makes sense to assume that organizations do have responsibilities of their own.

To live in a well-organized society with functioning corporate agents makes it easier for us, as individuals, to live up to our moral responsibilities. If the morally required task at hand is too big for us, individually, all we have to do is basically to dial the right number – call in the police, the ambulance, the firefighters, or whatever other corporate agent is suitable for the job. Thus the moral responsibilities for many of those actions which we can only perform jointly thus comes to be placed on the broad (if somewhat metaphorical) shoulders of our corporate agents.

Yet even in near ideal circumstances, this is not always enough. Even if the police, the ambulance etc. are well organized, they cannot be everywhere all the time. Thus it can be morally required of us, the individuals, to team up and organize ourselves quite spontaneously to do the right thing, at least until the relevant corporate agent appears on the scene – even if we're not friends who are already used to deliberate about what to do, how to do it, and to act together. In some situations, there is the moral responsibility to team up with strangers – people who happen to be around simply by chance. This is what has come to be discussed under the title "the responsibility of random collections".

What kind of responsibility is at stake here? As this seems to be somewhere between the two cases, one way to approach this question is from the distinction between the responsibility

of individuals and the responsibility of corporate agents. Thus conceived, the question is this: Is our duty to act jointly more like an extension of our individual responsibility to dial the right number to teaming up, or is it more like the responsibility of a make-shift or substitute "Citizen's Police's" responsibility to do their job? The difference is whether all responsibility involved is individual responsibility, or whether there is responsibility that is a collective's. The strongest intuitive argument against the latter view is that in order to have responsibility, there has to be an agent whose responsibility it is, and that while there are group agents, "random collections" are not of that sort. Or, put differently: for somebody or something to be morally responsible, that somebody or something must be, under suitable circumstances, an adequate target of moral blame; yet if people fail to team up to come to your aid, it might seem you would not rationally blame "the bystanders", collectively, but rather each of the individual bystanders, individually. Random collections, it seems, cannot be the target of moral blame, because they are no agents. The strongest intuitive argument against the first view is that where strangers fail to team up and do what's morally required of them to do, it seems what they are responsible for is the failure to act – and this is not exactly the same as the participant's failure to team up – in this sense, there seems to be a way in which such cases do involve a responsibility that is the collective's after all.

The main thrust in the received literature on the responsibility of random collections is to accommodate some of the latter intuition, while remaining firmly committed to the view that whatever responsibility random collections might have cannot be genuinely collective, but has to be some distribution of individual responsibility (I.). In this view, a bunch of strangers simply isn't the kind of entity that is a suitable target for blame. A random collection of strangers is no collective agent, especially where the random collection fails to team up. Yet I shall argue that the distributive view of the moral responsibility of random collections makes it hard to give an adequate understanding of how the responsibility in question is not just to team up, but rather to act, too. I will address an attempt in the recent literature to solve this dilemma by separating collective responsibility from collective agency (II.) and develop an alternative view in which there is a condition under which the collective responsibility of random collections is a plural agent's. Where this condition is met, random collection of individuals can be the kind of entity that is a suitable target for blame even where they failed to team up. The condition in question is the capacity for plural self-awareness, that is, the right sort of practical knowledge of what the strangers should be doing together (III.).

Co-subjective consciousness creates collectives

MICHAEL SCHMITZ
University of Vienna

ABSTRACT

Approaches to collective intentionality are commonly distinguished in terms of whether they locate collectivity in the content, mode or subject of intentional states, that is, in terms of what subjects believe or intend (e.g. Bratman), in a special we-mode (e.g. Searle, Tuomela), or in collective, plural subjects (e.g. Gilbert, Schmid). Despite their differences, all these approaches take for granted the standard model of intentional states as propositional attitudes.

On this model, which is inspired by reports of attitudes, the subject and the attitude type do not contribute to the intentional content of the attitude, which is exhausted by what the subject(s) perceive, do, believe, or intend. In my contribution I argue that we should rethink our understanding of propositional attitudes: the subject is not only aware of the state of affairs it believes to obtain or intends to bring about, but also has at least a sense of her own practical or theoretical position towards that state of affairs. For example, when it intends something, it must have a sense that it's up to itself to bring about the intended state of affairs, that it has taken up practical responsibility for doing so. So the subject not only represents a state of affairs, but its position vis-à-vis that state of affairs through what I call "position mode content" and itself through what I call "subject mode content" – since it can only represent its position if it represents itself.

I go on to argue that this revised understanding of intentional states, which can also be extended to speech acts – I use "postures" as a cover term for both – can be the basis for an improved understanding of collective intentionality. The key idea here is that when we take up the kind of postures that are constitutive for the existence of collectives, others are represented as co-subjects of these positions towards the world rather than as their objects. This is shown on three levels of collective intentionality. These levels are distinguished in terms of the format or structure of the relevant intentionality – the pre-conceptual level of joint attention and joint bodily action, the conceptual level of joint intention and belief and common knowledge, and the documental level of institutional reality, where we take up postures in institutional roles such as being a clerk, judge, or professor. For each level, I demonstrate the benefits of understanding collective intentionality in terms of co-subjective consciousness and subject mode representation.

Joint attention is a matter of attending with rather than to somebody else and thus distinct from mere mutual attention. This distinction can be explained in terms of how the co-attenders figures in the intentional content of their joint attention experiences: as co-subjects rather than as objects. This content is affectively charged and disposes to joint action. This PAIR-account of joint attention as a pragmatic and affectively charged intentional relation between co-subjects can explain a variety of results from developmental psychology. For example, infants understand which objects co-attenders have interacted with and which are new to them

before they understand such things with regard to others that they merely observe and thus experience as objects rather than as co-subjects (Moll & Meltzoff 2011). Autistic children show characteristic deficits in their affective and attentional relations to others and in co-subjective styles of reference to them.

On the basis of the PAIR account we can further explain the conceptual irreducibility of the “we” to the “I” which has been argued by Searle and others at the level of joint belief, intention and knowledge: “we” can’t be reduced to “I” because in the fundamental cases it refers to subjects as related to each other in the attentional, actional and affective ways described by the PAIR account. At the same time, it shows that there is nothing mysterious about this “we”. It is not free-floating with regard to the individuals, but just picks them out as related in a certain way – as co-subjects. Moreover, the subject mode account can easily avoid the infinite iteration of states that mar traditional approaches such as those of Lewis and Schiffer. The subjects of common knowledge don’t have beliefs of the form “I know that you know that p” and “You know that I know that you know that p”, and so on, but their relevant thoughts and beliefs are simply of the form “We know that p”, that is, they represent each other as co-subjects of knowledge.

With regard to institutional reality I argue that it is best understood in terms of what I call “role mode”, which is a form of subject mode. That is, in the fundamental case, institutional reality does not – counter to what notably Searle has argued – exist because of what we believe, desire or intend. For example, a professor is not a professor because others believe that she is one, but because she takes theoretical and practical positions from the vantage point of this role, and because others also represent and accept her in that role from the vantage points of their roles as her students, colleagues, administrators and so on, and by accepting the rights and obligations that come with the roles that are defined in relation to her role. Moreover, all have to represent themselves as co-subjects in an organization that has common purposes.

In conclusion, I argue that the co-subject approach overcomes the opposition between content, mode, and subject approaches to collective intentionality by extending content to mode and explaining collectives in terms of co-subjective consciousness and mode representation. In this way we can be unabashed realists about collective subjects while demystifying them at the same time.

The non-epistemic origins of research strongholds

MIKE D SCHNEIDER

UC Irvine

ABSTRACT

The goal of the present project is to investigate how homophily, the tendency for individuals who co-identify to preferentially associate with each other, manifests as a causal factor in the dispersion of new practices across diverse communities. Specifically of interest are scientific communities, i.e. communities whose very existence is predicated on the production of increasingly sophisticated knowledge claims about particular features of the world. These communities, by virtue of their epistemic commitments, manifest an external measure by which to study how social dynamics like those driven by homophily matter to the success of the community.

Pointing out that social identity matters in scientific communities is nothing new (e.g. Nelson (1993); Fehr (2011)), but so far there has been little explicit work toward identifying how it is that homophily can matter to the success of epistemic communities. I find that two distinct species of homophily that are present in the history of science can, in conjunction with one another, lead to the formation of research strongholds, preventing the spread of more appealing scientific practices from a minority population to the community at large.

First, I rehearse three historical cases from the literature to motivate a reading of homophily in scientific communities as being comprised of two distinct species. The first of these species of homophily, which I call sociological homophily, refers to the empirical observation that there tend to be fewer regular information channels across social identity groups than there are within social identity groups. The second species of homophily, which I call normative homophily, refers to the existence of a normative pressure to conform to the research choices of those within one's peer group. I argue that the conjunction of these two species of homophily has historically carried negative epistemic consequences, at least across the three cases considered.

Since historical investigations are largely silent on the subject of how it is that homophily can do what it seems to do, I then develop a formal model of homophily in scientific communities. In particular, I introduce a multi-type random network model of scientific collaborations that is sensitive to both species of homophily, in which agents play coordination games with each of their collaborators, i.e. each of their neighbors in the network. In the model, sociological homophily amongst researchers induces a particular sort of collaboration structure throughout the scientific community and, by virtue of this induced network structure, normative homophily is then capable of functioning as a causal mechanism affecting a variety of social epistemic phenomena that depend on the coordination of activities amongst collaborators.

Results of agent-based simulations of the model across a wide parameter space show how the presence of homophily in a scientific community gives rise to non-trivial epistemic consequences concerning the dispersion of research practices (and with those practices, the spread of beliefs and knowledge claims) throughout the community. In particular, I show how the

presence of homophily in a scientific community gives rise to segregated research strongholds, making it both less likely that the community converges to the more appealing of two choices of research practices and more likely that the community stays indefinitely divided. Although these results are unhappy when research strongholds are based on social identity, I conclude with a discussion of a slightly different model, in which the formation of research strongholds is epistemically motivated, in the spirit of preserving epistemic diversity (Strevens, 2003; Zollman, 2010). This leaves open the question of how it is that a particular scientific community might transform from one that features the unhappy research strongholds into one that features the epistemically defensible ones.

REFERENCES

- Fehr, Carla (2011). "What Is in It for Me? The Benefits of Diversity in Scientific Communities." *Feminist Epistemology and Philosophy of Science: Power in Knowledge*. Ed. E. Heidi Grasswick. Dordrecht: Springer Netherlands. 133– 155.
- Nelson, Lynn Hankinson (1993). "Epistemological communities." *Feminist epistemologies*, 83, 121–160.
- Strevens, Michael (2003). "The role of the priority rule in science." *The Journal of philosophy*, 100(2), 55–79.
- Zollman, Kevin JS (2010). "The epistemic benefit of transient diversity." *Erkenntnis*, 72(1), 17–35.

Joint and mutual obligations

DAVID SCHWEIKARD

Europa-Universität Flensburg

ABSTRACT

Suppose several agents have a joint obligation, i.e. an obligation to Φ that can only be fulfilled by means of a joint effort, then what are the corresponding obligations of the individual agents involved? In this paper, I take a closer look at how joint obligations are fulfilled and at the nature and extent of the corresponding individual obligations. As a first pass, one might say that joint obligations imply contributory obligations on the part of individual agents, so that each is under an obligation to do their bit of the required joint action. I shall argue that an adequate analysis will have to be more complex and encompass not just contributory obligations but a range of mutual obligations obtaining between the agents involved.

The starting points for my analyses will be variations on a case in which an agent finds herself in a dangerous situation and is in need of others' help. In particular, I shall take up the following scenarios:

(1) While out on a swim in a lake, agent A has suffered a cramp and cannot reach the shore by her own effort. She audibly cries for help. B is the only agent nearby who could rescue A. To do so, B would have to push a rowing boat into the water, row to A's position and rescue her.

(2) While out on a swim in a lake, agent A has suffered a cramp and cannot reach the shore by her own effort. She audibly cries for help. B and C are both near enough to rescue A, and each of them finds herself at a spot on the shore where there happens to be a suitable rowing boat. Each of them, B and C, could perform the rescue effort described in (1).

(3) While out on a swim in a lake, agent A has suffered a cramp and cannot reach the shore by her own effort. She audibly cries for help. Agents D and E are nearby and together find themselves at a spot on the shore where there happens to be a rowing boat that they can only move and navigate together. Thus they could only perform the rescue effort together.

For the following I assume that obligations obtain in all these scenarios: an individual obligation to help in (1), several individual obligations to help in (2), and a joint obligation to help in (3). (I here use "obligation" throughout, setting aside the question whether "duty" may be more appropriate.) I first suggest an analysis of the fulfilment of an individual obligation to help. Secondly, I discuss how the normative situations of the potential helpers in (2) change in comparison to (1). This will require distinguishing between the case in which B and C are unaware of the other's presence, and the case in which they are aware of each other. Thirdly, I analyse the fulfilment of the joint obligation obtaining in (3) with special focus on the individual and mutual obligations of the potential helpers D and E.

1. Individual obligations and their fulfilment

B's fulfilment of her obligation to help described in scenario (1) may include the following elements: she would need to (i) realise that A needs help, (ii) realise that she (B) could provide this help, (iii) recognise that she (B) should provide this help (or has an obligation to help), (iv) decide to help A, (v) perform the rescue effort. This list is not exhaustive but takes up relevant

elements that will prove useful in the discussion of the fulfilment of joint obligations. In particular, I don't claim that all these are necessary elements of the fulfilment of an obligation. By the term "realise" in (i) and (ii) I mean that B needs to become aware of the situation and form specific judgements. This invokes specific epistemic requirements to draw certain inferences concerning A's situation and B's own capacity for action. Clause (iii) requires that B recognise her obligation to help and thus form a specific normative judgement in the situation. Setting aside B's motivational situation, (iv) and (v) point out that the fulfilment of the obligation requires both decision and action.

2. Isomorphic individual obligations

In the variation of scenario (2) in which the two agents are unaware of each other, their respective normative situations are the same as in the case with only one potential helper present. Here it needs to be discussed how it can be prevented that they block each other's performance of their rescue efforts.

With respect to the variation in which they are aware of each other, their respective normative situations don't change fundamentally. In particular – as I will argue – their individual obligations to help are not weaker or (only) conditional just because there is another agent with an isomorphic obligation. A need for coordination may arise on the condition that only one can perform the complete rescue effort. This need highlights a first form of mutual obligations to coordinate; it may be that one of them acquires an obligation of omission in order to prevent a collision of their efforts.

3. Joint obligations and corresponding individual obligations

The conditions of fulfilling the joint obligation in (3) can be modelled on those specified for the individual case, but they are more complex insofar as the agents involved need to form not only specific individual attitudes but also joint attitudes. These are, in particular, joint beliefs regarding A's situation, their joint capacity for action and their joint obligation, as well as a joint decision and the intentional performance of a joint action.

Since ex hypothesi D and E can only perform the rescue effort jointly and have an irreducibly joint obligation, none of them has precisely this obligation to help. As I will seek to show, they instead each have obligations regarding the coming about of the joint attitudes mentioned and the performance of the joint action.

Grounded Composition?

JEROEN SMID

Lund University

ABSTRACT

Composite objects are very intimately related to their parts. For example, wherever the whole is, its parts are there, too. One may want to explain why this is the case. One explanation could be that composition is (a kind of) identity relation. This is the (strong) Composition as Identity thesis, which makes mereological composition equivalent to singular-plural identity. More recently, Ross Cameron (2014) suggested an alternative explanation based on the notion of grounding. Roughly, the suggestion is that a whole and the relation obtaining between the parts and the whole are grounded in the existence of the parts. Such theories are of interest to social ontology, because many social objects seem to be composite. I argue, however, that the grounding approach fails to explain the intimate relation between a whole and its parts.

REFERENCE

Cameron, Ross. "Parts Generate the Whole, but They Are Not Identical to It." In *Composition as Identity*, edited by Donald L. M. Baxter and Aaron J. Cotnoir, 90–110. Oxford and New York: Oxford University Press, 2014.

Artificial agents in the realm of social cognition

ANNA STRASSER

Humboldt-Universität zu Berlin

ABSTRACT

A growing number of philosophers are concerned about the problem that standard philosophical notions have in capturing critical socio-cognitive phenomena in other disciplines. Consequently, minimal versions of these standard notions were developed to capture a wider range of socio-cognitive abilities (cp. Butterfill and Apperly 2013; Michael et al. 2016; Vesper et al. 2010).

As artificial agents are increasingly prevalent in human social life (e.g., care robots, chat bots), it is important to consider how to capture our relations to them. We are still on the threshold of developing an appropriate conceptual framework for describing human-computer interaction in detail. By evaluating how minimal approaches can facilitate explaining socio-cognitive abilities in artificial agents I will contribute to this debate. Focusing on joint actions I discuss a critical approach toward the standard notion of action, as well as what role social cues and potential ascriptions of commitment have in interactions with artificial agents.

For this aim I introduce a minimal notion of action which is applicable to artificial systems and enables a finer-grained characterization of different types of actions. Thereby, I draw a new line between action and mere behaviour. Focusing on joint actions, I ask whether and under what circumstances artificial agents can count as proper agents. If artificial agents convincingly play a role in joint actions, we might stop to categorize them as mere tools. Research pertaining to conversational agents aims to develop artificial agents from mere tools into human-like partners (Mattar & Wachsmuth 2014). Social cues play a major role in joint actions. To better distinguish joint-actions from tool-use I examine reciprocal exchanges of social cues in human-computer-interactions. Evaluating to what extent multimodal social interactions can be handled, actual research will be examined (e.g. Kang et al. 2012; Petta et al. 2011; Becker & Wachsmuth 2006; Baur et al. 2015).

One perspective is to evaluate how the interaction with such agents is experienced by humans (Biocca et al. 2001; Bailenson et al. 2001). But the central issue here is not whether agents successfully ‘trick’ humans into ascribing properties to them that they do not have; rather, the issue is to develop a conceptual framework to characterize social abilities they actually do have, and to articulate benchmarks for roboticists in refining those abilities further. The 3D humanoid agent MAX, a conversational machine developed by the research group of Ipke Wachsmuth (Bielefeld), is an excellent candidate to serve as a possible test case. Besides the ability to interpret and produce social cues, commitment plays an important role in joint actions. Therefore, it is important to discuss whether we may interpret behaviour of artificial agents as signs of ‘commitment’. Thanks to a minimal approach introduced by Michael et al. (2016), we can address the question as to what kinds of commitments towards artificial agents might be engendered. According to this minimal approach several components of a

strict commitment can be dissociated and, consequently, the single occurrence of one component is treated as a sufficient condition for a minimal sense of commitment. Those minimal cases are of special interest regarding artificial systems. Applying this minimal structure to a human-computer-interaction leads to controversial claims. It exhibits cases in which a human is committed toward an artificial agent, as well as cases in which humans expect the artificial agent to be committed. Conversely, it characterizes cases in which the artificial system is committed to the human, or expects the human to be committed. Discussing those cases will illuminate difficult and controversial questions. Up to now we have had no clear-cut conditions regarding the ascription of commitment to artificial systems. But, on the other hand, conversational agents such as MAX display behaviours seemingly of being offended in reaction to unfulfilled 'expectations'. It is uncontroversial that humans ascribe mental states to non-living beings, and this can support the success of social interactions. Building upon this, this paper investigates whether a sense of commitment may arise in such interactions, what form it may take, and what role(s) it may play.

REFERENCES

- Bailenson, J. et al. (2001). Intelligent Agents Who Wear Your Face: Users Reactions to the Virtual Self. In: Antonio, Aylett, Ballin (eds.) *Intelligent Virtual Agents*, 86. Springer: Berlin.
- Baur, T. et al. (2015). Context-Aware Automated Analysis and Annotation of Social Human-Agent Interactions. *ACM Transactions on Interactive Intelligent Systems* 5.2.
- Becker, C. & Wachsmuth, I. (2006). Modeling primary and secondary emotions for a believable communication agent. In: Reichardt, Levi, Meyer (eds.) *Proceedings of the 1st Workshop on Emotion and Computing*.
- Biocca, F., Harms, C., Gregg, J. (2001). The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. *4th annual International Workshop on Presence. Philadelphia, PA*.
- Butterfill, S. & Apperly, I. (2013). How to Construct a Minimal Theory of Mind. *Mind and Language* 28 (5) 606-637.
- Kang, S. et al. (2012). Towards building a Virtual Counselor: Modeling Nonverbal Behavior during Intimate Self-Disclosure. *Eleventh International Conference on Autonomous Agents and Multiagent Systems. Valencia, Spain*.
- Mattar N. & Wachsmuth, I. (2014). Let's get personal: Assessing the impact of personal information in human-agent conversations. In: Kurosu (ed.) *Human-Computer Interaction. Advanced Interaction Modalities and Techniques*. Lecture Notes in Computer Science. Vol 8511. Berlin: Springer.
- Michael, J., Sebanz, N. & Knoblich, G. (2016). The Sense of Commitment: A Minimal Approach. *Front. Psychol.* 6:1968.
- Petta P., Pelachaud C., Cowie, R. (eds.) (2011). *Emotion-Oriented Systems: The Humaine Handbook*. Heidelberg/Dordrecht/London/New York: Springer.
- Vesper, C., Butterfill, S., Sebanz, N. & Knoblich, G. (2010). A minimal architecture for joint action. *Neural Networks*, 23(8/9), 998–1003.

Nested Groups: Membership and Parthood

DAVID STROHMAIER

University of Sheffield

ABSTRACT

Recently, the ontology of groups has received increasing attention within philosophy of the social sciences (Uzquiano 2004, Ritchie 2013, 2015, Epstein forthcoming). The question typically asked regards the reducibility of groups to other kinds of entities—i.e., sets or fusions. Here I seek to address two issues of group ontology often overlooked in the literature.

One issue is that many groups are nested (see Thomasson forthcoming). Nested social groups bear specific relations to one another such that, at least metaphorically, one group is within another. My aim is to clarify what precisely the nesting relation is and I will argue that there are in fact two nesting relations: membership and parthood.

The second issue relates to the scope of group membership—in particular, whether only individuals can be members. I demonstrate that our most plausible account of nested groups leads to an expansion of the scope of membership.

Clarifying such issues of group ontology allows us to consider more complex candidates for groups, such as states. Such cases are difficult to analyse without looking at how the group relates to other groups. How does the United Kingdom, for example, relate to the Ministry of Defence? Is this, metaphysically speaking, the same relation the UK has to NATO? And who qualifies for membership in a group like the United Kingdom anyway?

In the remainder of this abstract I sketch the two kinds of nesting, why they are necessary, and how my argument affects the scope of membership.

Here is an example: A marketing department is nested in the hypothetical corporation CORP. I call a lower-level group, like CORP's marketing department, a "sub-group", and a higher-level group, like CORP, a "super-group". In addition CORP belongs to a cartel that fixes prices. Relative to the cartel, CORP is a sub-group and relative to CORP the cartel is a super-group. Three nested levels of groups are at work here.

There are relations which the marketing department, CORP, and the cartel bear to one another which make them sub- and super-groups. In what follows, I call a relation, which makes two groups nested, the nesting relation between sub- and super-group. I propose that two different nesting relations hold, the first being mereological parthood, i.e., the marketing department is a part of CORP. The second relations is that of group membership, i.e., CORP is a member of the cartel.

The above conflicts with a popular conception of membership, which constrains membership to individuals (List & Pettit 2011, Effingham 2010: 259). In this view, the relation between sub- and super-groups is that of parthood, membership being limited to individuals. The UK would be – metaphysically speaking – a part of the NATO, rather than a member.

I argue against this view by showing that there is a difference between two types of nesting-cases. There is one class of nested groups where, necessarily, if x is a member of the sub-group, x is a member of the super-group. There exists, however, another class, where it is not the case that, necessarily, if x is a member of the sub-group, x is a member of the super-group. In

other words, the classes of nested groups differ in whether sub-group members are necessarily super-group members or not. Here is an example for the two classes: A member of CORP's marketing departments, let's call him Henry, is automatically a member of CORP itself. However, a member in CORP is not automatically a members of the cartel. This difference, I suggest, is best explained by parthood being the nesting relation between the groups in one case and membership in the other. In this way, I show that limiting nesting relations to parthood is a mistake and we should broaden the scope of membership so that the UK can be a member of NATO. To show the details of my argument let me look at the example of Henry a bit more. Because Henry is a member of CORP's marketing department, he has to be a member of CORP. There is no way around that. I suggest as an explanation that the nesting relation between marketing department and CORP ensures that every member of the sub-group is a member of the super-group. Let's call the nesting relation R-lower.

There is also a higher-level nesting relation between CORP and the cartel, which I call R-higher. Can R-lower and R-higher be the same relation? No, they cannot and for the following reason: Henry is a member of CORP, but by all appearance he is not a member of the cartel. If the cartel members were convicted of cartel activity, then Henry would not be convicted as one of them. He might have some responsibility for CORP being a cartel member, but he is not a member himself. So R-higher does not ensure that the members of the sub-group are members of the super-group. But R-lower ensured just that! It follows, that R-lower and R-higher must be different relations. Accordingly, the view that all sub-groups stand to the super-group in the relation of parthood fails. My own view that nesting relations can be parthood and membership relations remains as the better solution.

I therefore conclude that we should adapt our analysis of the membership relation to allow it to hold between groups so as to account for nested groups. Such an account of nested groups can make sense of the difference between R-lower and R-higher, in contrast to accounts that limit the nesting relation to parthood. This result is a step forward in analysing complex groups: The UK can be a member of NATO, with all that entails.

Collective Hatred and Hatred towards Collectives

THOMAS SZANTO

University of Copenhagen

ABSTRACT

Hatred figures prominently among negative collective affects. Indeed, among 'affects of aggression' (Demmerling & Landweer 2007), hatred occupies not only a particularly extreme but, moreover, a structurally distinctive place. I will argue that hatred is an essentially collective affect or emotional disposition. Specifically, I shall claim that affective-intentional acts realizing this disposition can only be directed towards social facts, groups or (alleged) proxies of such. My claim, however, will be only modestly 'collectivist' in spirit, as I do not wish to imply that hatred can only be realized or enacted by groups as such.

In defending my modestly 'collectivist' account of hatred, I will draw upon two strands of research: on the one hand, on relevant proposals within the recent discussion on collective and political emotions (Schmid 2009, 2014; Nussbaum 2013; von Scheve & Salmela 2014; Sanchez Guerrero 2016); on the other hand, on two congenial phenomenological conceptions proposed by Kolnai (1932) and Ahmed (2004), both of which conceive of hatred as an essentially social or collective phenomenon.

I will present my argument in three steps: (1) First, I will explore the affective-intentional structure of acts of hatred and differentiate between its intentional target or object, and its affective focus (Helm 2007). I will show that the intentional target of hatred is always a person or groups of persons (and never an ordinary material object), and that unlike, for example, in (collective) fear, the affective focus of hatred is not some potentially eliminable property of the respective person or group, nor the causal effects they may have upon oneself or one's group (danger, harm, etc.). Rather, it is some negatively appraised socio-cultural norm or value instantiated by the target, or, as Kolnai (1932) views it, the 'existential' role upon one's 'way of life'. Consequently, I shall argue that targets are typically actual or alleged representatives of certain groups or certain ideologies, religions, or socio-cultural norms. Moreover, I shall argue that the affective-intentional focus of hatred is typically 'diffuse', oscillating between concrete individuals and representatives of groups. Thus, the negatively evaluated characteristics of the hated person(s) are typically attributed to representatives of 'imagined communities'. (2) In a further step, I will examine the dialectical relationship between the subject of acts of hate and the hated (personal) object with reference to Ahmed's conception of a certain 'affective alignment' and Kolnai's 'existential unity' between addressors and targets. In particular, I will explore the dialectic of exclusion or elimination of the hated group or representatives, and a positive affective re-affirmation of or alignment with the 'community of haters'. (3) Finally, I will explore the socio-psychological dynamics (e.g. Alford 2011) underlying this dialectic, the role of hatred for social identification and in-group/out-group demarcations, and their 'sedimentation' in something like a 'habitus of hatred' (Kolnai 1932).

Exploitation and Joint Action

ANDRÁS SZIGETI AND ERIK MALMQVIST

University of Linköping

ABSTRACT

This paper aims to forge a link between exploitation and collective action theory. We do this by focusing on a kind of exploitation that has remained largely unexamined until now despite the growing interest in the concept of exploitation and exploitative practices.

The concept of exploitation is attracting renewed attention in moral and political philosophy, partly due to a growing philosophical interest in allegedly exploitative practices such as commercial surrogacy, sweatshop labour, and medical research in the developing world. Yet exploitation theory remains underdeveloped in an important way. The philosophical debate has concentrated on what might be called *direct* exploitation which occurs when one party to an exchange or relationship (*A*) takes advantage of the other party (*B*) in order to benefit himself/herself. At the same time, philosophers have for the most part ignored what might be called *indirect* exploitation. This occurs when *A* takes advantage of *B*, but the gains *A* extracts from *B* accrue not (only) to *A* but (also) to *C*. This omission is striking because indirect exploitation appears to be widespread in the real world. For instance, sweatshops may exploit workers in poor countries on behalf of multinational firms and for the benefit of their customers. Similarly, clinical researchers and their sponsors may exploit research subjects to advance the interests of future patients.

Our work takes up such cases of indirect exploitation. We distinguish between different kinds of indirect exploitation depending on the type of involvement by the third party. We also argue that different types of involvement have different moral implications especially as regards the remedial duties incurred by the third parties. Our main claim is that in many cases third-party beneficiaries can be properly construed as *acting together* with exploiters in bringing about and/or maintaining the exploitative relationship or transaction. Establishing joint action in such cases is important because, first, it helps to explain why third parties are morally responsible for the exploitation. Second, the third parties' moral responsibility in turn provides a rationale for allocating duties to remediate the victims of exploitation to them.

The paper starts by briefly characterizing exploitation in the standard two-party sense, and by outlining the debate about its moral wrongness. We then clarify what we mean by direct exploitation as opposed to indirect exploitation and using simplified cases distinguish three main ways in which third-party beneficiaries can be involved in indirect exploitation: (i) through joint action with the exploiter; (ii) culpably, but without joint action; and (iii) non-culpably benefiting from exploitation.

We provide a detailed conceptual analysis of (i), i.e., indirect exploitation by joint action. We discuss the relevance of the threefold distinction, and defend it against objections. We argue that our taxonomy corresponds to differences in third-party duties of remediation towards victims of exploitation.

We close by summing up the main findings and consider the possibility of real-life applications. Consider, for instance, a multinational firm that does business with a developing world

subcontractor that exploits its workers. Should the firm be classified as innocently benefiting from the exploitation, as knowingly contributing to it, or as acting jointly with the exploiter – or perhaps as an exploiter, period? And how should we classify an infertile Western couple that seeks out a surrogate mother in India with the assistance of an exploitative local reproductive clinic? How we answer these questions bears on the moral assessment of the firm's or the couple's conduct, and specifically, on how one characterizes their duties to the exploited workers or the exploited surrogate.

Needless to say, determining the nature of the third party's involvement in each particular real life case will be heavily dependent on a wide range of facts on the ground. In order to know whether a third-party innocently or culpably benefits from exploitation, and whether it cooperates or not with exploiters, we have to find out what the third party knows, what its intentions are and how these intentions relate to those of the exploiter, and what courses of action were available to it under the circumstances. In short, our taxonomy of indirect exploitation offers no shortcut to evaluating indirect exploitation in the real world because adequately analyzing such cases will require coupling our framework with careful attention to empirical detail. However, we hope that our taxonomy highlights what *kind* of features to look for when we attempt such analyses.

Is Art an Institution? Are Artworks Social Objects? Does the Philosophy of Art Fit into Social Ontology?

ENRICO TERRONE

University of Turin

ABSTRACT

According to institutional accounts of art (cf. Dickie 1969, Iseminger 2004, Abell 2012), art is an institution within which artworks are produced and appreciated. Yet, if one conceives of an institution as a system of rules that governs social interactions (cf. Guala 2016, 3-19), thereby comparing art with fully-fledged institutions such as money or marriage, one finds it hard to figure out the system of rules that should constitute art as a unitary institution (cf. Lopes 2014, 108-115).

I will argue that art is not an institution but rather a practice, understood as a network of convergent attitudes and basic, often unstated, presuppositions (cf. Brandom 1994, 21; Lamarque 2010, 10). While institutions are basically made of rules, practices basically make rules. While art as an institution should be a *system of rules* within which artworks are produced and appreciated, art as a practice constitutes each artwork by bestowing a *specific rule* on it.

More specifically, art as a practice constitutes artworks in conformity with Searle's (1995) formula for the construction of social reality: the object X counts as having the status function Y in the context C. In the case of art, C is the practice itself, X is an object embedded within this practice, and Y is a status function that prescribes to evaluate X according to what I call a *Criterion of Appreciation* (henceforth, CoA).

The CoA establishes what features are relevant to the appraisal of X. Thus, Y, as a deontic function, provides X with *the power* to be assessed according to its CoA: a beholder *ought to* evaluate X according to its CoA, at least if she wants to count as a suitable appreciator of X.

The notion of a CoA has been prefigured by Neo-Aristotelian metaphysicians like Strawson and Wiggins, who conceive of the artwork as a special individual, which requires a competent attitude on the part of the subjects that evaluate it. In Strawson's terms, "The criterion of identity of a work of art is the totality of features which are relevant to its aesthetic appraisal" (1966, 202). In Wiggins' terms: "We think of the artist as making something which is calculated by him to have an effect which cannot be characterized in instrumental or non-aesthetic terms, and cannot be identified independently of some totality of relevant features" (2001, 137). A similar point has been made by Irvin: "When an artist puts forward an object with certain features, he or she is sanctioning the set of *artwork* features that, given the context and the conventions connecting the object and the artwork, the suitably informed audience will take the artwork to have" (2005, 322).

From this perspective, art involves a game of justified appraisal whose rules are specific for each artwork on which a CoA has been bestowed. Art, so understood, is not a *system* of rules, namely an institution, but rather a practice that allows us to establish a (potentially unlimited)

series of rules, that is to say, a CoA for every artwork.

‘Do you like it? Why?’ are the basic questions that a beholder faces when confronted with an artwork. In principle, the beholder can answer whatever she wants, but in justifying her answer she should make reference to the features established by the CoA. Imagine two persons debating about a work of art. One says ‘I like this work because p ’ and the other replies ‘No, I don’t like it, because q ’. The CoA of that artwork specifies the features to which propositions such as p and q can correctly make reference.

This leads us to highlight the key difference between artworks and other objects of *aesthetic appraisal* as for example flowers or landscapes (by ‘aesthetic appraisal’ here I mean an evaluation that derives from a subjective state of pleasure or displeasure and makes claim to correctness by attributing to the object a value that depends on the features of this very object, cf. Zangwill 2014). The aesthetic appreciation of artworks is socially governed in a way in which the appreciation of other objects is not. An artwork, unlike a flower or a landscape, has a CoA that establishes the totality of features that are relevant to its aesthetic appraisal.

If all of this is right, we can finally address the three questions raised by the title of this paper. Is art an institution? No, it is rather a practice. Are artworks social objects? Yes, since they are provided with a distinctive status function, which prescribes us to evaluate each artwork in accordance with its own CoA. Does the philosophy of art fits into social ontology? Yes, provided that we acknowledge that artworks are special social objects in that each of them has a specific status function, which accurately prescribes how to evaluate it. While the institution of money entitles us to evaluate a variety of things in the same way, the practice of art commits us to evaluate each artwork in accordance with its own CoA. If one takes money as the paradigmatic social fact, art seems to lie at the opposite end of the spectrum. But the end of the spectrum is still a part of it.

REFERENCES

- Abell, C. (2012), “Art: What it Is and Why it Matters”, *Philosophy and Phenomenological Research*.
- Brandom, R. (1994), *Making it Explicit*, Harvard University Press.
- Dickie, G. (1969), “Defining art”, *American Philosophical Quarterly*.
- Guala, F. (2016), *Understanding Institutions*, Princeton University Press.
- Irvin, S. (2005), “The Artist’s Sanction in Contemporary Art”, *The Journal of Aesthetics and Art Criticism*.
- Iseminger, G. (2004), *The Aesthetic Function of Art*, Cornell University Press.
- Lamarque, P. (2010), *Work and Object*, OUP. Lopes, D. (2014), *Beyond Art*, OUP.
- Searle, J. (1995), *The Construction of Social Reality*, Free Press.
- Strawson, P. (1966), ‘Aesthetic Appraisal and Works of Art’, *The Oxford Review*.
- Wiggins, D. (2001), *Sameness and Substance Renewed*, CUP.
- Zangwill, N. (2014) ‘Aesthetic Judgment’, in *The Stanford Encyclopedia of Philosophy*.

Two Ways of Feeling Together

GERHARD THONHAUSER

Free University Berlin and University of Vienna

ABSTRACT

Within the debate on the nature of collective intentionality, it is often assumed that there is one prototypical notion of collectivity, which needs to be traced back to some intentional feature of an attitude, whether this feature is its subject, mode, or content. Discussing the case of shared emotions, I argue that we need to distinguish several senses in which an attitude can be shared. In my talk, I focus on two senses of feeling together, which I label emotional collectivity and emotional sameness.

Emotional collectivity signifies affective attitudes that are phenomenologically and functionally ours: we experience them as our emotions, and they are also constitutively not mine and your, but ours. Emotional collectivity involves some form of self-awareness as a group and presupposes some processes of social cognition. These kinds of collective attitudes have been discussed extensively in the debate on collective emotions (cf. Gilbert 2002; 2014; Helm 2008; 2014; Schmid 2009; 2014; Salmela 2012; Salmela and Scheve 2014; Sánchez Guerrero 2016). Whereas there is disagreement as to where the collectivity of the collective emotion should be located, most agree that it needs to be based on some coherent set of features that apply to all cases. There is debate on how much diversity among the feelings of the involved individuals is feasible in emotional collectivity, and some claim that we need to account for various types of sharing. However, these debates are still based on the idea of one prototypical notion of collectivity.

To develop an understanding of emotional sameness, I build on the early phenomenologist Max Scheler. Scheler (1913) introduced a typology of ways in which we relate to the emotional attitudes of others. To begin with, Scheler uses the category of ‘sensing’ (Nachfühlen), which is roughly similar to what other phenomenologists discussed under the label of empathy and what current debate investigates as social cognition. Scheler underscores that when sensing the emotion of another, I am not experiencing the same or a similar emotion. I rather perceive the emotion as the emotion of the other, without partaking in the foreign attitude. Second, Scheler defines ‘fellow feeling’ (Mitgefühl) as sensing plus an emotional reaction to the other’s emotion through an emotional attitude of one’s own. It is important to notice that in fellow feeling, the emotion of the other and my emotion are two separate emotions – no actual sharing is involved. Scheler draws a sharp distinction between these two forms of social cognition and forms of sharing. Moreover, he introduces two distinct forms of sharing, claiming that one is prior to sensing and fellow feeling, whereas the other is only possible on their basis. He characterizes the first as a form of ‘feeling-the-same’ on the basis of ‘emotional contagion’ (Gefühlsansteckung), whereas he labels the latter as ‘feeling-with-one-another’ or ‘co-feeling’ (Miteinanderfühlen). Co-feeling is analog to what is currently discussed under the labels shared or collective emotion. It is hence not surprising that co-feeling has been the focal point of the recent rediscovery of Scheler (Schmid 2009, Krebs 2010, Salice 2016, Szanto 2016). Feeling-the-same, on the other hand, signifies a form of sharedness that is characterized by a lack of

awareness of the participant's individuality. It neither presupposes an awareness of the plural nature of the attitude nor of each other's participation (Schlossberger 2016).

Emotional sameness is a form of feeling together based on emotional contagion that does not fit with the prototypical notion of collective attitudes dominant in current debate on collective emotions. However, I suggest that it can be brought in fruitful dialog with psychological and sociological research on the manifold perspectives in which emotional sharing can be found (Scheve and Ismer 2013), especially concerning the primary modes of interaction and the micro-level mechanisms of emotional convergence. Although it is difficult to find any strict criteria to distinguish emotional sameness and emotional collectivity, it is reasonable to say that emotional collectivity is linked to processes of social cognition and group identification, and involves patterns of social appraisal and a reference to social practices, norms, and values. Emotional sameness, on the other hand, is based on emotional contagion and a direct resonance of expressive behavior. In this regard, Scheler's analysis of feeling-the-same has a lot in common with traditional crowd psychology (Le Bon 1896, Freud 1921, Reich 1933). In light of current political developments, it appears urgent to study such contagious mechanisms of emotional sameness that fall beneath the standards of rationality and awareness that are usually the threshold for philosophical account of collective intentionality.

Collectivity in Blockchain based Environments

SABINE THUERMEI

Munich Center of Technology in Society, TU Muenchen

ABSTRACT

In the last years a novel variant of acting together came into being. It is still in its embryonic stage. It is characterized by trustless interactions between humans who have interdependent interests respectively share goals. Their unique way to deal with uncertainties regarding each other is the “binding of the will by code and to the corresponding technology platforms”. The most prominent example is the blockchain technology. Non-trusting blockchain users can interact with each other without a trusted human intermediary. Blockchain technology serves as the basis to digital currencies as Bitcoin, to a wide variety of distributed ledgers, where cryptographic tokens represent ownership of material goods as diamonds or land (smart property), and to self-executing smart contracts. In the future decentralized autonomous organizations (DAO) may be launched based on smart contracts and code. Their autonomy results from the fact that after the instantiation they no longer need their creators. Moreover, they are self-sufficient in so far that they can accumulate capital both digital currencies and physical assets. Their business model consists in charging for digital services provided. As long as they have sufficient funds they can operate in an independent way. An ill-intentioned DAO may cause irreparable havoc.

The legal consequences of smart contracts, where and under what circumstances they have legally binding effects, as well as approaches to distributed governance of such platforms are currently hotly discussed subjects: Instead of relying on trust between humans the blockchain based distributed databases intend to provide trustworthy, i.e. incorruptible records of business transactions. Instead of relying on human institutions to enforce contracts contractual partners may rely on block-chain based platforms for self-executing smart contracts. Instead of trusting central banks users of digital currencies trust free market mechanisms supported by blockchain technology. Since technology may also not be faultless these technologies do not result in a trust-free world but in one where we trust technology more than each other. Trust is put solely in nonhumans potentially resulting in a trustless society. Executing a smart contract is decoupled from concluding such a contract between humans or technical agents. As the adherents of the Lex Cryptographia say: “code is law”. In these environments individual freedom might be maximized possibly to the detriment of others. However, the old ways to collaborate as plural subjects, in we-mode, in modest sociality continue to exist. Responsible innovators start to think about how the opportunities provided by blockchain technologies may be preserved while distributed meritocratic governance models and alternative dispute resolute methods may be incorporated in blockchain platforms. So even if today (collective and shared) responsibility is completely delegated to the technology focusing on a few norms as incorruptibility this need not be the case in more mature approaches.

Collective testimony and second personal commitment

LEO TOWNSEND

Centre for the Study of Mind in Nature

ABSTRACT

In this paper I offer a novel account of collective testimony, according to which the speech act of collective testimony involves a certain kind of collective epistemic commitment. More specifically, the speech act of a group telling an audience that *p* is one that involves the group second-personally committing to *p*'s being true. I develop my proposal against the backdrop of another, somewhat similar approach to collective testimony that has recently been suggested by Miranda Fricker (2012). According to Fricker, a group's testimony that *p* involves a joint commitment, on the part of group's members, to epistemic trustworthiness with respect to the matter of *p*.

The first part of the paper is an exposition and critique of Fricker's account. At the heart of her account is the notion of 'deal' of second-personal epistemic trust that gets struck between the parties in a testimonial exchange. The speaker's side of this deal involves her making a commitment to prove trustworthy with respect to the matter about which she speaks, while the hearer's side involves a reciprocal commitment, the commitment to trust the speaker in what she says. Fricker attempts to extend this general approach to testimony from the individual to the collective case by appropriating Margaret Gilbert's concept of 'joint commitment' (see, e.g., Gilbert 2014), but I argue that this is an unfortunate, yet telling, choice. This is because, firstly, Gilbertian joint commitments are ill-suited to the epistemic domain, and, secondly, they cannot adequately accommodate the normative import of seeing testimony in second-personal terms.

In the second part of the paper I present an alternative account to Fricker's, which aims to retain the core insights of her approach, while avoiding its shortcomings. In short, I think that Fricker is right that collective testimony involves the group making a kind of epistemic commitment, and she is right that testimony must be understood as an essentially second personal I-thou affair. But where she goes wrong, I contend, is in how she connects these two features in her account. Specifically, she locates the second-personal character of testimony in the object of the group members' joint commitment – i.e. in what they are jointly committed to, namely, to proving to be second-personally trustworthy in some epistemic matter. My proposal, by contrast, is that the object of the commitment the group makes in testifying is simply the truth of what is attested, and the second-personal nature of the commitment comes not from this object but from the kind of commitment it is, i.e. a second-personal commitment.

I thus set about developing my alternative proposal in two main stages: first, clarifying the nature of second-personal epistemic commitment, and, second, explaining how a group can successfully commit itself in this distinctive way. In the first stage I draw on recent work by Richard Moran (2013), which connects the Searlean tradition of thinking about assertives in terms of the speaker's overt epistemic commitment with an account of illocutionary speech acts that sees them as fundamentally second-personal. According to Moran, such acts 'are

themselves essentially relational between one person and another, [...] which establish or alter the normative relations between them' (Moran 2013: 133). In the second stage I draw on recent work by Philip Pettit (2014, 2016) which focuses on the role played by authorized corporate spokespersons in enabling groups to speak for, and not merely about, themselves.

WORKS CITED

- Fricker, M., (2012). 'Group Testimony? The Making of A Collective Good Informant,' *Philosophy and Phenomenological Research* 84, pp. 249–276.
- Gilbert, M., (2014). *Joint Commitment: How We Make the Social World*. Oxford University Press: Oxford.
- Moran, R., (2013). 'Testimony, Illocution and the Second Person'. *Aristotelian Society Supplementary Volume*. 87, pp. 115–135.
- Pettit, P., (2014) 'How to tell is a group is an agent' in Lackey, J (ed.) *Essays in Collective Epistemology*. Oxford University Press: Oxford.
- Pettit, P., (2016) 'The Birth of Ethics' in Mark Matheson (ed.) *The Tanner Lectures on Human Values*, Vol. 34, University of Utah Press: Salt Lake City.

On the moral normativity of socially constructed norms

LAURA VALENTINI
LSE

ABSTRACT

Socially constructed (or “positive”) norms—namely the formal and informal norms that regulate our interpersonal, social, and political lives—are all around us: from the British practice of queuing, to the Western custom that we shake someone’s hand when we meet them for the first time; from the norm that one should dress modestly in church, to the complex demands that entire legal systems place on us.

Socially constructed norms present a puzzle for moral theorizing. On the one hand, we feel their moral pull: we often believe that we would act wrongly if we were to violate their demands—at least when those demands are consistent with fundamental moral principles. On the other hand, it is not clear where that pull comes from. Positive norms are social phenomena. They exist by virtue of people’s collective attitudes: their believing in certain oughts and their being disposed to act on them, as a matter of common knowledge (e.g., Brennan et al. 2013; Searle 2006). The puzzle consists in explaining how the fact that people have certain beliefs and dispositions (i.e., an “is”) gives rise to moral obligations to act in particular ways (i.e., “oughts”).

Two solutions to this puzzle are particularly prominent. Neither of them, I think, is fully satisfactory. The first, which I call the deflationary view, consists in denying that positive norms have any moral normativity (cf. Enoch 2011). On this view, whenever we think that we ought to do as those norms require, this is by virtue of the contingent merits of the norms’ prescriptions: the way they further some independent good. So, for example, we should queue at bus stops because doing so is an effective coordination mechanism; we should wear modest clothes in Church because doing otherwise would upset some people (and it is bad to upset people), and so forth. The difficulty with the deflationary view is its inability to account for a class of intuitive instances of wrongdoing: those that involve harmless breaches of positive norms. Consider, for instance, harmless breaches of promises and cases of harmless trespass. While such actions strike many of us as wrongful, their wrongness cannot be traced to their setting back some independent good. On the deflationary view, these wrongs must be assumed away, and this seems unsatisfactory.

This difficulty with the deflationary view does not affect the normative interests view, defended by David Owens (Owens 2012; forthcoming). This view expands the realm of value, by postulating a class of interests that Owens calls “normative.” On Owens’ view, for instance, we have an interest in having the power to control others’ obligations (an “authority interest”), and this explains why even harmless violations of promises wrong us: they violate this authority interest. This is only one type of normative interest, to which Owens adds several others: deontic, permissive, and remissive interests. By postulating the existence of this class of interests, the normative interests view can explain instances of wrongdoing that the deflationary view must assume away. This strength of the normative interests view, however, is also its weakness.

Postulating normative interests is unparsimonious, and potentially ad hoc; the sort of move we should make only if nothing else could account for our intuitive judgements about wrongdoing. My aim in this paper is to show that we need not choose between (i) explaining away our intuitive judgements and (ii) unparsimoniously postulating new entities to make sense of them. Instead, I show that a familiar moral principle—which I call the “agency respect” principle—grounds the moral normativity of socially constructed norms. To do so, I distinguish—in line with insights from the Kantian tradition (e.g., Noggle 1999; Hill 2000)—between requirements of person respect and of agency respect. Requirements of person respect specify how we ought to treat individuals given their status as autonomous end-setters in general, independently of their particular ends and commitments. They set out minimal moral standards applying to our treatment of persons *qua* persons (e.g., standards forbidding certain forms of harm and mandating certain forms of assistance). Requirements of agency respect, by contrast, tell us how we ought to treat others given the ways in which they exercise their agency, namely given their particular commitments, projects and pursuits. Agency respect gives us *pro tanto* obligations to accommodate those commitments and pursuits, provided they are consistent with person respect.

Armed with the agency-respect principle, it is easy to see where the moral normativity of socially constructed norms comes from. This is because socially constructed norms are underpinned by agential commitments. They depend, ontologically, on particular exercises of people’s agency. Recall, for a positive norm to exist in a certain context, a large enough number of people must believe in the content of the norm (the relevant “ought”) and be disposed to act on it, as a matter of common knowledge. If the content of the norms under discussion is consistent with person respect, the obligation to agency-respect those whose agential commitments sustain them (“norm-supporters”) infuses those norms with moral normativity. This explains where the moral normativity of positive norms comes from. The wrong of harmless breaches of positive norms—including of promising norms, property norms, courtesy norms, and others—stems from the fact that such breaches violate the *pro tanto* obligation to agency-respect norm-supporters. This, I conclude, offers a parsimonious and satisfactory solution to the puzzle of the moral normativity of socially constructed norms.

The Plurality in a Plural Subject: The Phenomena of Dissenting and Dissidence as Communal Phenomena

XIAOXI WU

University of Vienna

ABSTRACT

People act together, they share plans and are engaged in common endeavors. It intrigues philosophers to inquire what exactly such “togetherness” amounts to. One important effort of spelling out what constitutes togetherness is in Searle’s seminal work, where he calls it “a sense of us” (1990, 145). This sounds intuitively appealing, for we do feel a sense of being together, “a sense of us”, when we are engaged in collective activities. Schmid (2014) takes up the concept of “a sense of us” and develops a plural subject account of collective intentionality. Schmid argues that the sense of us is actually plural pre-reflective self-awareness, which constitutes plural subject. Plural self-awareness is fundamental and does not arise in virtue of communication and interaction between individual subjects.

In a series of papers Zahavi and his co-authors (2015, 2016, 2017) raise worries about Schmid’s account. The central point is that granting the plural subject such fundamentality seems to fail to capture the plurality of the participants that make up the plural subject. After all, sharing can only occur between individual subjects. In my talk, I shall argue that Schmid actually has the resource to accommodate the requisite plurality, because Schmid (2009, 47-58) takes note of the phenomena of dissenting and dissidence, which are the strongest case that exhibits the plurality of relevant participants. Furthermore, the sort of mutualist accounts along the line Zahavi advocates seems to fall short of accounting for the phenomena of dissenting and dissidence, which presupposes a primordial pre-reflective commonality despite of the disagreements concerning what “we” should do on the reflective level.

The idea of Schmid’s account is to take the pre-reflective self-awareness at the individual level to the collective level, arguing that the first-personal character of self-awareness does not have to be singular. It can also be plural, i.e. first-person plural. Since the basic traits of singular self-awareness are ownership, perspective, and commitment, correspondingly the basic traits of plural self-awareness are common ownership, shared perspective, and joint commitment. These traits should not be understood as the “properties” of a pre-existing plural subject. Instead, they constitute the plural subject and are what it means to feel “a sense of us”, to be a collective. And such “sense of us” is prior to the differentiation of self and other, and thus not achieved from such differentiation through interaction. According to Schmid, attempts to account for sharedness by virtue of communication seem to suffer from circularity. Because, in order for any communication to take place, some sharedness has already been in place.

Zahavi challenges the priority of the plural self-awareness and argues that the self-other distinction is actually prior to and persevered in the plural subject, “sharing has nothing to do with fusion, nor with a merged unity, sharing involves a plurality of subjects, but it also

involves more than mere summation or aggregation.” (2015, 92) He argues for the priority of self-other differentiation drawing upon resources from both classical phenomenology and developmental psychology, and adopts a sort of mutualist account of collectivity intentionality, namely, the sharedness is achieved in and through second-person perspective taking, which roughly amounts to being engaged in reciprocal encounters between individuals. Individuals communicate and interact, a “we” thereby arises out of such reciprocal activities. Thus, on Zahavi’s account, we start with a plurality of individuals, and achieve the unity of a collective through communication, in contrast to Schmid’s account according to which communication is only possible when we are already plurally self-aware.

However, I would like to argue that without some prior pre-reflective sense of us that is the very foundation of communication, it seems difficult for Zahavi’s mutualist account to accommodate the phenomena of dissenting and dissidence, which play important roles in the dynamics of collective practice and communal life. (Sustein 2003)

Zahavi endorses Szanto’s (2015) proposed requirements for an account of sharing to be viable, i.e. the plurality condition and the integrity condition. Intuitively, sharing is between various individuals. So, the very concept of sharing presupposes a plurality of participants. However, communication alone is not enough to secure the integrity condition. For instance, when two people are engaged in heated argument against each other, they are obviously communicating and interacting. But, Zahavi argues, since the confrontational aspect is too salient, the integrity condition, which requires a sort of unity of a collective, is not met. Thus, the sort of communication Zahavi elaborates seems to be quite agreeable. It has to be so, for there is no other way to achieve the requisite unity of a collective than reaching some sort of agreement or amicable atmosphere in and through communication. Then, it is hard to see how Zahavi can accommodate dissenting. A dissenting person or group has a sharply different view concerning what we should do together than the established one. The tension between the dissenters and non-dissenters seems to be stronger and more global than the tension between two people who happen to be involved in a heated argument.

In contrast, Schmid has the advantage of easily accommodating the highly confrontational phenomena of dissenting and dissidence, without thereby losing the unity that brings a collective together as a whole. For, on his account, the requisite unity does not depend upon the agreement people happen to reach. The plural self-awareness that constitutes the collectivity lies in the pre-reflective level, it is the common ground that makes it possible that members of a community can have heated disagreements over how their shared life should be led and how they should make progress together. These disagreements are rather different articulations and developments of what is already shared, they stimulate a healthy dynamic that makes the unity of the communal life be able to ever renew itself without thereby suffering from disintegration. Thus, Schmid’s plural self-awareness account not only respects the plurality of the participants in a collective endeavor, but also brings the deep root that grounds such plurality to the fore.

REFERENCES

- Schmid, H. B. (2009): *Plural Action*. (Dordrecht: Springer)
- (2014): "Plural Self-awareness," *Phenomenology and the Cognitive Sciences* 13: 7-24
- Searle, J. (1990): Collective Intention and Action, in *Intention in Communication*, ed. Pollack et al (MIT Press), 404-15
- Sustein, C.R. (2003): *Why Societies Need Dissent*. (Harvard University Press)
- Szanto, T (2015): "Husserl on Collective Intentionality," in *Social Reality: The Phenomenological Approach*, ed. Salice and Schmid. Dordrecht: Springer
- Zahavi, D (2015): "You, Me, and We: The Sharing of Emotional Experiences", in *Journal of Consciousness Study*, Volume 22, No.1-2, 84-101
- with Brinck and Reddy (2016): "The Primacy of the 'We'?", in *Embodiment, Enaction, and Culture*, ed. Durch, Fuchs, and Tewes, (MIT Press), 131-47
- (2017): "Second-person Engagement, Self-alienation, and Group-identification," in *Topoi*, 1-10, forthcoming

What would you like me to drink?

The Epistemic Foundations of Salience-based Coordination

VOJTĚCH ZACHNÍK

University of Ostrava

ABSTRACT

This paper aims to assess current theoretical findings on the origin of coordination by salience and suggests a way to unify the existing framework. The main concern is to reveal how different coordination mechanisms coexist in a mutually consistent way by pointing out specific epistemic aspects. The paper highlights the fact that basic epistemic assumptions of theories diverge in a way that makes them not only mutually compatible, but also interlinked. Consequently, recommendations and predictions of the unified view of coordination are, in principle, based on the processes related to the agent's presumptions regarding the cognitive abilities of a co-player.

There are many coordination challenges in our everyday lives: how to cross the road safely, which greeting pattern to choose, who goes through the door first, and how to arrange a meeting. Yet, we do not feel that these kinds of everyday interactions involve some sort of obstacle, because our behaviour usually seems straightforward and effortless. The key puzzle that will accompany us in the paper is how this behaviour emerges the first time, and that we subsequently feel secure in continuing it. What are the underlying processes enabling this type of interdependent behaviour of multiple agents? How many different but parallel ways can bring about coordination in general? The notion of *salience* was preliminarily specified in terms of "standing out" or "conspicuousness" (Schelling 1960; Lewis 1969), which broadly explains its role in the decision-making process. Namely, an individual who wants to coordinate with others, but does not know which behavioural pattern is precisely suitable for the situation, may look to the assistance of a salient feature of an interaction (contextual clue, labelling of choice, etc.).

In general, my aim is to assess recent leading proposals and answer the following question: Is it the case that people coordinate by salience because they frame contextual cues and conceive the situation from a new perspective (Bacharach and Bernasconi 1997; Bacharach and Stahl 2000; Bacharach 2006), or because individuals have a "hunch" about another's behaviour and try to respond to this as best as possible (Camerer, Ho, and Chong 2004)? I can reveal in advance that my answer is 'yes' in both cases, since both views are built upon some elementary assumptions about the beliefs of others, and the logical structure of these epistemic foundations ensures that these two approaches are compatible. Therefore, I suggest there are parallel kinds of salience-based coordination processes, and that their usage is determined by the epistemic context of an interaction. More specifically, the focus of interest is a particular model of interactions and the role of salience in these interactions. The phenomenon of salience is placed in a more refined theoretical foundation that provides more robust explanatory grounds for both interpreting experimental data, and reflection on different modes of reasoning. I

will briefly introduce two prominent theories of salience-based coordination – *variable frame theory*, and *cognitive hierarchy theory* – and then reveal their epistemic background to prove their mutual compatibility.

The argument runs as follows: first, components of theories exhibit remarkable differences regarding two aspects – correct beliefs and beliefs in rationality. Hence, even before it comes to establishing coordination every strategically thinking agent makes some estimates concerning possible interaction scenario, and co-player's behaviour and beliefs. Second, it is suggested to bring into play terms of *epistemic symmetry* and *asymmetry* and to argue that they aptly elucidate process of selection when one of the coordination procedures is initiated instead of the other. As a result of this distinction, a coordination game with salience allows a number of diverse but parallel procedures. Either an agent assumes that he and his co-player are symmetric in important epistemic aspects, or he expects asymmetric conditions to be valid. In the first situation, correct beliefs, and belief in rationality, are prerequisites for the use of subsequent framing and the application of team reasoning. Whereas in the second, the epistemic type of agents is such that they rather anticipate some level of incorrectness in beliefs and uneven standards of rationality, which leads to the utilization of best-response reasoning, based on each agent's cognitive efforts.

Therefore, an implication of such reflections regards the relationship between symmetrical and asymmetrical perspective. I consider them as mutually exclusive alternatives: each time one is well-founded and dominant, the other is absent and ruled out. Based on these grounds, it is more than plausible to say that variable frame and cognitive hierarchy theory are like two sides of the same coin. Under suitable epistemic circumstances, it is likely that an agent is well equipped to use the respective coordinating principle determined by the theory without ruling out future use of another principle.

In sum, coordination by salience is a result of the involvement of two processes whose active impact on the final outcome is fundamentally – but not solely – determined by the epistemic niche of a given interaction. And if one wants to properly understand to coordination by salience it seems necessary to take into account the epistemic restrictions that are imposed on reasoning when individuals follow different coordination procedures. Two relevant and prominent theories, variable frame theory and cognitive hierarchy theory, should be regarded as complementary modes supporting the same goal of coordinating individuals' actions. And suggested criteria of epistemic symmetry and asymmetry comprehensively specify when and how such coordination emerges.

Belief-Revision, Epistemic Contribution, and Polynormativity

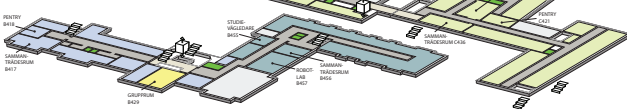
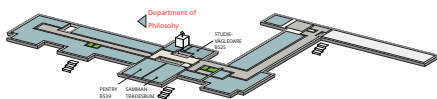
MARKO-LUKA ZUBČIĆ

Faculty of Humanities and Social Sciences in Rijeka

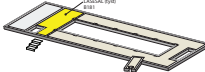
ABSTRACT

Given the limited inferential capacity of any human epistemic agent, the best social epistemic system includes as many human epistemic agents as possible and has them “working under” diverse epistemic norms. In this text, this claim is argued for through presenting a pragmatist and instrumentalist argument for Epistemic Contribution and, consequently, the diversity of epistemic norms (polynormativity). It is argued that 1) the design of the social epistemic system hinges primarily on the chance of the revision of false belief, and 2) through universal inclusion and polynormativity we raise our chances of the revision of false belief. Furthermore, showing how neither Dewey’s democracy nor Hayek’s markets can by themselves sustain not slipping into epistemically distortive social arrangements, I argue, along Mill, that there should be an institutional order that primarily maintains universal inclusion in and polynormativity of epistemic cooperation. Certain tentative requirements of this institutional order are discussed through focus on issues of the development of diverse normative communities and zones of their engagement, understanding epistemic suboptimality, freedom of exit, and epistemically distortive inequalities.

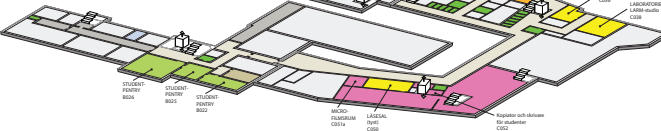
Maps



B339 GRUPP-Raum B339 B336



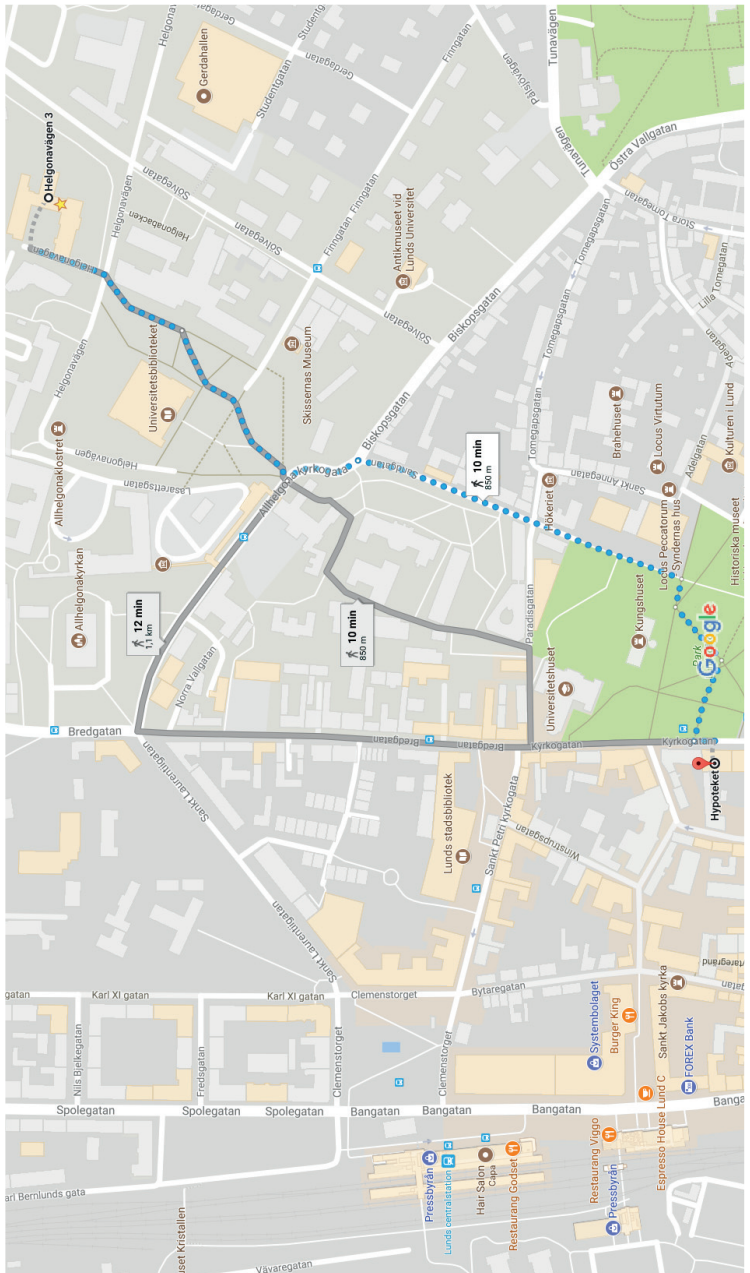
LUX-biblioteka



- [illegible]

Google Maps Helgonavägen 3, Lund to Hypoteket, Kyrkogatan, Lund
LUX Building to Hypoteket (venue for conference dinner on Thursday)

Walk 850 m, 10 min



Google Maps LUX Building to ENSO V lunch place, Restaurang Valvet



Restaurang Valvet
Allhelgona kyrkogata 7, 223 62 Lund

Kartdata ©2017 Google

Program

Day 1 (Wednesday Aug 30th)

8.00 - 9.00	Registration (by C121)
9.00 - 9.30	Welcome (C121)
9.30 - 10.45	Keynote: Karen McComb, <i>Sociality, communication and cognition in non-human animals</i> Chair: Tomas Persson (C121)

Coffee (by C121)

	Chair: Olle Blomberg (B240)	Chair: Anna-Sofia Maurin (B336)	Chair: Megan Lambert (B339)	Chair: Andras Szigeti (B352)
11.05 - 11.35	Leo Townsend <i>Collective testimony and second personal commitment</i>	Arto Laitinen <i>Declarations and Institutional Reality</i>	Sabine Thuermel <i>Collectivity in Blockchain based Environments</i>	Jules Salomone <i>From Coordination to Cooperation. Switching to Team Reasoning</i>
11.40 - 12.10	Kourken Michaelian <i>Collective memory: Metaphor or reality?</i>	Tobias Hansson Wahlberg <i>Causal Powers and Social Ontology</i>	Sigmund Schilpzand <div>CANCELLED</div>	Hein Duijf <i>Collective reasoning and collective responsibility gaps</i>
12.15 - 12.45	Pekka Mäkelä, Raul Hakli and S.M. Amadae <i>Institutions as Shared Plans: Practical Reasoning in I and We Modes</i>	Benjamin Neeser <i>Objects, People, and Powers: The Deontic Theory of Money</i>	Anna Strasser <i>Artificial agents in the realm of social cognition</i>	

13.00 Lunch at Valvet

	Chair: Raul Hakli (B240)	Chair: Jeroen Smid (B336)	Chair: Gunnar Björnsson (B339)	Chair: Åsa Burman (B352)
14.00 - 14.30	Michael Schmitz <i>Co-subjective consciousness creates collectives</i>	Esa Diaz Leon <i>Social Construction and Universality</i>	Gerhard Thonhauser <i>Two Ways of Feeling Together</i>	Joshua Rust <i>Money and Deontic Power</i>
14.35 - 15.05	Glenda Lucila Satne <i>Non-observational knowledge and Joint action</i>	Robin Guid Loher <i>How to Think About Constructions: Towards an Account of the Acquisition of Social Kind Concepts</i>	Antonella Carassa and Marco Colombetti <i>Normative Representations and Emotions</i>	Gloria Mähringer <i>Collective Intentionality as the common ground of normative experiences</i>
15.10 - 15.40	Anna Moltchanova <i>A theory of collective self-awareness</i>	Frank Hindriks <i>What's the Difference Between Money and Gender? Social Construction, Critique, and Change</i>	Matteo Bianchin <i>Social Cognition, Cooperation, and Hypothetical Contracts</i>	Hannah Rubin and Cailin O'Connor <i>Discrimination and Collaboration in Social Networks</i>

Coffee (by C121)

16.00 - 18.00	Symposium: <i>From Individual to Plural Agency</i> Kirk Ludwig, Sara Rachel Chant, Michael Schmitz, Raimo Tuomela, Chair: Björn Petersson (C121)
---------------	--

Reception

Day 2 (Thursday Aug 31st)

9.00 - 10.15 Keynote: Christian List, *Levels: Descriptive, Explanatory, and Ontological*
Chair: András Szigeti (C121)

Coffee (by C 121)

	Chairs: Michael Schmitz / Julie Zahle (B240)	Chair: Tobias Hansson Wahlberg (B336)	Chair: Frank Hindriks (B339)	Chairs: Stephen Butterfill / Thomas Szanto (B352)
10.35 - 11.05	Judith Martens (Chair: Michael Schmitz) <i>Heuristics, bounded rationality, and joint action</i>		Gunnar Björnsson <i>Moral demands on groups</i>	Vojtěch Zachník (Chair: Stephen Butterfill) <i>What would you like me to drink? Epistemic Foundations of Salience-based Coordination</i>
11.10 - 11.40	Olle Blomberg (Chair: Michael Schmitz) <i>More (unsound) arguments for the common knowledge condition</i>	Francesco Franda <i>Organizations: An Ontological Approach</i>	Hans Bernhard Schmid <i>Collective Responsibilities of Random Collections: Plural Self-Awareness Among Strangers</i>	Mike D Schneider (Chair: Stephen Butterfill) <i>The non-epistemic origins of research strongholds</i>
11.45 - 12.15	Taina Kalliokoski (Chair: Julie Zahle) <i>The Explanatory Power of the Axiological Objects of Collective Intention</i>	David Strohmaier <i>Nested Groups: Membership and Parthood</i>	Eirik Julius Risberg <i>Collective Agency and Individual Criminal Responsibility</i>	Kalle Grill (Chair: Thomas Szanto) <i>A New Method for Social Preference Aggregation</i>
12.20 - 12.50	Abe Roth (Chair: Julie Zahle) <i>The Practical Significance of Group Reasons for Individual Agents</i>	Jeroen Smid <i>Grounded Composition?</i>		Bo Allesøe Christensen (Chair: Thomas Szanto) <i>Taking one for the team. Team-reasoning and the possibility of collective alief.</i>

13.00 Lunch at Valvet

14.00 -	Keynote: Sara Rachel Chant, <i>From Minimal to Complex Collective Actions</i>
15.15	Chair: Gunnar Björnsson (C121)

Coffee (by C 121)

	Chair: Glenda Satne (B240)	Chair: Jeroen Smid (B336)	Chair: Arto Laitinen (B339)	Chair: David Schweikard (Note room! B 237)
15.45 - 16.15	Brian Flanagan <i>The Theory of Group Agency: A Synthesis and an Experiment</i>	Gustav Ramström <i>Macro Phenomena without Macro Manifestations – Why the Multiple realization Argument is Incompati- ble with Social Science</i>	Mattias Gunnemyr <i>Explaining Moral Responsibility for Structurally Produced Harms</i>	Ingvar Johansson <i>How Do Non-Joint Commitments Come Into Being? : An Attempt at Cultural Naturalism</i>
16.20 - 16.50	Thomas Szanto <i>Collective Hatred and Hatred towards Collectives</i>	Gianluca Pozzoni <i>Social Reality Without Levels</i>	Anton Eriksson <i>Causal Responsibility for Emissions in Global Supply Chains</i>	Henning Nörenberg <i>Nostra Res Agitur? Pro- grammatic Contexts and The Disentangled We</i>
17.00 - 18.15	Keynote: Stephen Butterfill, <i>Cooperation and Motor Representation</i> Chair: Olle Blomberg (C121)			

Conference dinner

Day 3 (Friday Sep 1st)

	Chair: Tomas Persson (B240)	Chair: Anna-Sofia Maurin (B336)	Chair: Gunnar Björnsson (B339)	Chair: Johan Brännmark (B352)
9.00 - 9.30		Meghan Masto <i>CANCELLED</i>	David Schweikard <i>Joint and mutual obligations</i>	Eyja Brynjarsdóttir <i>Money as an Interactive and Subjective Social Kind</i>
9.35 - 10.05	Xiaoxi Wu <i>The Plurality in a Plural Subject: The Phenomena of Dissenting and Dissidence as Communal Phenomena</i>	Rachel Fraser <i>Towards a Social Ontology of Work</i>	András Szigeti and Erik Malmqvist <i>Exploitation and joint action</i>	Åsa Burman <i>The ontology of human rights and obligations</i>

Coffee (by C121)

	Chair: Elisabeth Pacherie (B240)	Chair: Tobias Hansson Wahlberg (B336)	Chair: Mattias Gunnemyr (B339)	Chair: Eyja Brynjarsdóttir (B352)
10.25 - 10.55	Ingar Brinck <i>Minimal approaches to joint action</i>	Enrico Terrone <i>Is Art an Institution? Are Artworks Social Objects? Does the Philosophy of Art Fit into Social Ontology?</i>	Stephanie Collins and Holly Lawford-Smith <i>When the (Democratic) State Acts, does its Citizenry?</i>	Marko-Luka Zubčić <i>Belief-Revision, Epistemic Contribution, and Polynormativity</i>
11.00 - 11.30	Anika Fiebich <i>Social Cognition, Empathy, and Agent-Specificities in Cooperation</i>	Alessandro De Cesaris <i>A Ghost Story. Toward an Hegelian Approach to Social Ontology</i>	Niels de Haan <i>Collective Moral Agency and Identity Over Time</i>	Laura Valentini <i>On the moral normativity of socially constructed norms</i>
11.45 - 12.45	ISOS assembly meeting (C121)			

13.00 Lunch at Valvet

14.00 - 15.15	Keynote: Julie Zahle, <i>Explanatory and Ontological Levels in the Individualism-Holism Debate</i> Chair: Johan Brännmark (C121)
---------------	--

Coffee (by C121)

15.35 - 17.35	Symposium: <i>Understanding Institutions</i> Francesco Guala, Rico Hauswald, Elisabeth Pacherie, Wlodek Rabinowicz Chair: Åsa Burman (C121)
17.35 - 17.50	Thank you

